a folha

Boletim da língua portuguesa nas instituições europeias

http://ec.europa.eu/translation/portuguese/magazine

N.º 54 — verão de 2017

Software livre e de código aberto: OmegaT na DGT, Wizard, Tagwipe e TeamBase (parte 1)

Maria José Bellino Machado; Hilário Leal Fontes
Thomas Cordonnier; Elio Fedele; Fons De Vuyst
Direção-Geral da Tradução — Comissão Europeia

[versão inglesa deste texto — http://ec.europa.eu/translation/portuguese/magazine/documents/folha54_ot_en.pdf]

1. *Introdução*

Num artigo anterior em «a folha», publicado em 2014(1), a utilização de *software* livre e de código aberto na Direção-Geral da Tradução (DGT) — nomeadamente da ferramenta de tradução assistida por computador OmegaT(2) — foi apresentada em termos gerais. No presente artigo, que está dividido em duas partes, são apresentadas as seguintes aplicações para a tradução assistida por computador, recentemente disponibilizadas pela DGT da Comissão Europeia como *software* livre e de código aberto:

- **OmegaT na DGT** (DGT-OmegaT) com adaptações/melhorias/novas funções na implementação na DGT do OmegaT (versão *standard* 3.6(3)), desenvolvido internamente para satisfazer as suas necessidades (programador: Thomas Cordonnier)
- **DGT-OmegaT Wizard** (programador: Elio Fedele)
- **Tagwipe** (programadores: Elio Fedele e João Rosas)
- **Teambase** (programador: Thomas Cordonnier).

O projeto é gerido por Fons De Vuyst, chefe do Setor de Apoio Operacional da Unidade de Informática da DGT.

Na primeira parte deste artigo, apresentamos uma breve descrição de: a) fluxo de trabalho da DGT em cujo contexto foram adaptadas/desenvolvidas estas aplicações; b) Tagwipe; e c) 10 funções adaptadas

ou acrescentadas ao OmegaT. Na segunda parte deste artigo, que será publicada no final de 2017, apresentaremos o DGT-OmegaT Wizard e a TeamBase, duas aplicações que ainda exigem algum trabalho para serem «utilizáveis» fora da DGT.

Quanto à publicação do DGT-OmegaT como *software* livre e de código aberto, gostaríamos de sublinhar que o objetivo não é criar uma nova comunidade «em concorrência» com a equipa que desenvolve e mantém o OmegaT original. A DGT publica o código apenas para permitir a qualquer pessoa ver as funções que criámos para satisfazer as nossas necessidades, deixando ao critério da comunidade OmegaT decidir se algumas delas são de interesse geral e podem ser reutilizadas e integradas.

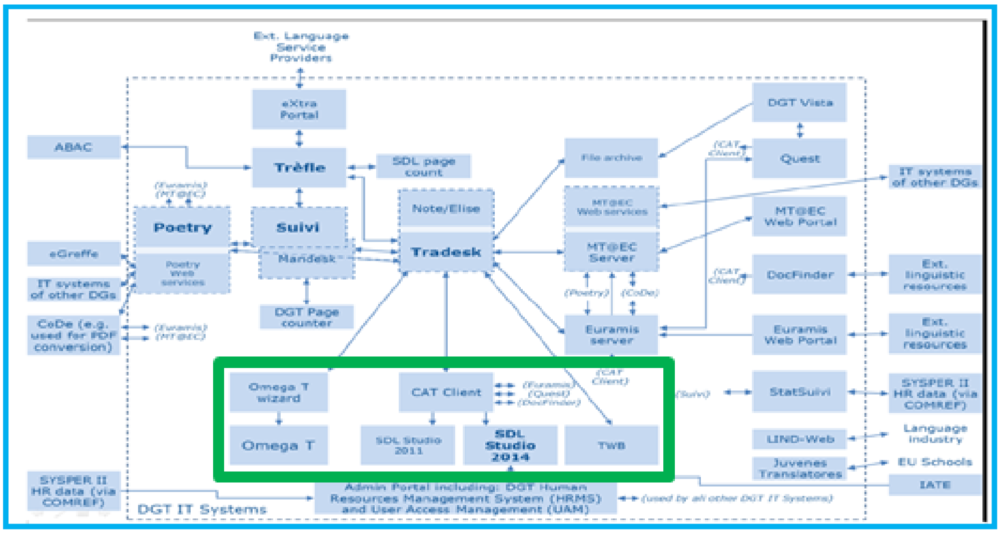
Estas aplicações podem ser descarregadas e facilmente experimentadas/utilizadas por qualquer pessoa em qualquer plataforma compatível com Java8 (Windows, Linux, MacOS), uma vez que são disponibilizados tanto o código-fonte como a versão executável, mas salientamos que são publicados tal como estão, sem garantias expressas nem implícitas, e que não será prestado qualquer apoio.

2. **Software** *livre e de código aberto na Comissão Europeia*

A Comissão Europeia tem vindo a promover a utilização — e a partilha — de *software* livre e de código aberto, designadamente no âmbito do 7.º Programa-Quadro de Investigação e Desenvolvimento Tecnológico(4) e, atualmente, do Horizonte 2020(5). Além disso, a Estratégia de *Software* de Código Aberto da Comissão Europeia(6) salienta que «os serviços da Comissão irão participar cada vez mais nas comunidades de *software* de código aberto a fim de desenvolver as aplicações de base de código aberto que são utilizadas no *software* da Comissão».

Por conseguinte, é neste contexto que as aplicações de código aberto como o OmegaT estão a ser utilizadas na DGT, paralelamente a aplicações comerciais como o SDL Trados Studio, a principal ferramenta de tradução assistida por computador na DGT.

A DGT tem utilizado o OmegaT para fins de prototipagem desde 2012 e, para esse efeito, a versão 2.6.0._3 do OmegaT foi inicialmente adaptada e complementada com algumas melhorias úteis. Alguns tradutores gostaram tanto do OmegaT que este foi incluído no **panorama informático da DGT** e, por essa razão, foi desenvolvido um **assistente** — o DGT-OmegaT Wizard — a fim de integrar o DGT-OmegaT no fluxo de trabalho específico da DGT.

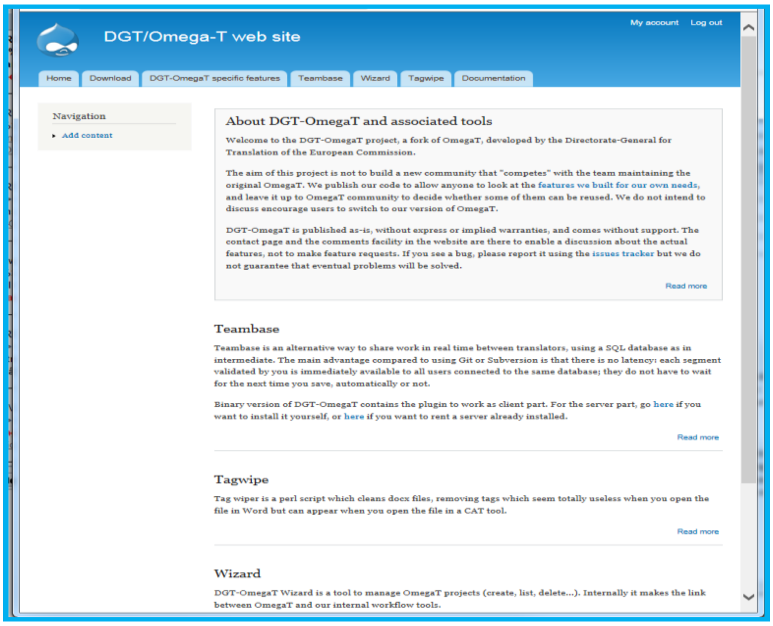


Panorama Informático da DGT
Direção Geral da Tradução, *Information Technology Master Plan 2016–2020*

Além disso, foram desenvolvidas duas outras aplicações que podem ser usadas tanto por utilizadores do OmegaT como do SDL Studio:

- **TagWipe**, uma aplicação que elimina etiquetas de formatação redundantes nos ficheiros em formato DOCX (Microsoft Office).
- **TeamBase**, que permite a partilha de memórias de projeto entre tradutores que podem utilizar uma ou outra das ferramentas de tradução assistida por computador disponíveis na DGT.

Recentemente, a DGT disponibilizou estas aplicações como *software* livre e de código aberto e as mesmas podem ser descarregadas aqui: http://185.13.37.79/.



3. *Fluxo de trabalho da DGT*

O ambiente de tradução assistida por computador da DGT é explicado em pormenor na publicação da DGT *Translation tools and workflow*(7). De salientar, para este artigo, que na DGT:

- **As memórias de tradução** relevantes para o(s) documento(s) a traduzir são extraídas automaticamente do Euramis, o repositório interno onde são guardados todos os alinhamentos das traduções da DGT — bem como de outras instituições da UE — nas últimas duas décadas e a maior parte da legislação da UE em vigor, em todas as línguas oficiais.
  Os segmentos relevantes e os documentos de referência são extraídos em formato TMX e utilizados como memórias externas numa das ferramentas. No DGT-OmegaT, estas memórias são automaticamente copiadas para a subpasta **\tm** de cada projeto pelo DGT-OmegaT Wizard.
- **A tradução automática** é gerada automaticamente para todos os documentos editáveis a traduzir. A tradução é feita pelo serviço MT@EC(8), o sistema de tradução automática interno da DGT baseado no sistema livre e de código aberto Moses(9), que está treinado com os *corpora* da UE e adaptado às necessidades da DGT.
  Os ficheiros TMX gerados pelo MT@EC são utilizados como memórias externas em ambas as ferramentas. No DGT-OmegaT, estas memórias são automaticamente copiadas para a subpasta **\mt** de cada projeto pelo DGT-OmegaT Wizard.
- **A terminologia** encontra-se guardada na base IATE, a base de dados terminológica interinstitucional da UE.
  Nos projetos DGT-OmegaT, os termos relevantes para o(s) documento(s) a traduzir são extraídos em formato TXT a partir de uma exportação (apenas termos de partida e de chegada) de toda a base IATE e o glossário «filtrado» é automaticamente copiado para a subpasta **\glossary** de cada projeto pelo DGT-OmegaT Wizard como um glossário só de leitura.

4. *Tagwipe*

Com relativa frequência, os documentos em formato DOCX (a esmagadora maioria dos documentos traduzidos na DGT) apresentam (muitas) etiquetas de formatação inúteis. Nesses casos seria obviamente impossível utilizar o OmegaT sem a função **Remove tags** ativada num número (infelizmente não tão raro quanto isso) de documentos particularmente «maus». No entanto, esta não é a melhor solução por razões óbvias.

Many point to possible further improvements in the use of the **ESIFs** and call for NGOs and local authorities to be given direct access to funds, for better enforcement of *ex ante* conditionalities, sanctions for failure to uphold the partnership principle, better monitoring (through an increased role for the Commission and Roma themselves) and action to prevent the ineffective use of funds (e.g. training programmes not leading to employment) or their misuse (e.g. ESIF interventions financing segregated settings), including through a transparent complaint mechanism.	<t76/>Many <t77/>point to<t78/> <t79/>possible further improvements <t80/>in<t81/> <t82/>the<t83/> use of<t84/> <t85/>the <t86/>E<t87/>SIF<t88/>s<t89/> and call for NGOs and local authorities<t90/> <t91/>to be given<t92/> direct access to funds<t93/>, <t94/>for <t95/>better enforcement of <t96/>ex ante <t97/>conditionalities, sanctions for <t98/>failure to uphold <t99/>the partnership principle, better monitoring (through<t100/> an<t101/> increased role <t102/>for <t103/>the Commission and Roma themselves) <t104/>and <t105/>action to <t106/>prevent<t107/> the ineffective use <t108/>of funds <t109/>(e.g.<t110/> <t111/>training programmes not leading to employment) <t112/>or<t113/> <t114/>their <t115/>misuse <t116/>(e.g.<t117/> <t118/>ESIF interventions financing segregated settings)<t119/>, including through<t120/> a transparent complaint mechanism<t121/>.

Por conseguinte, a DGT desenvolveu internamente a aplicação Tagwipe que elimina praticamente todas as etiquetas de formatação redundantes em documentos em formato DOCX. Melhora igualmente a segmentação das frases, algo que é muito importante para assegurar a coerência com a segmentação do Euramis e da MT@EC, melhorando assim a percentagem de correspondência (total ou parcial) com as memórias externas.

Na DGT, o Tagwipe é utilizado desde 2012, num ambiente Windows 7, para projetos DGT-OmegaT e, por conseguinte, é atualmente uma aplicação bastante estável. No sítio Web do DGT-OmegaT, é também disponibilizado o Tagwipe para a distribuição Linux Ubuntu.

No Tagwipe, o nível de limpeza das etiquetas de formatação pode ser escolhido pelo utilizador entre o nível 0 e o nível 8. Por omissão, na pasta de instalação, o nível de limpeza definido é o segundo nível mais baixo e mais conservador («**level.1**»). Este é o nível utilizado na DGT uma vez que os documentos traduzidos têm de ser gerados com toda a formatação original.

No entanto, quando não é esse o caso, o nível de limpeza pode ser aumentado, eliminando assim formatação não essencial como, por exemplo, cores de fundo de palavras/expressões, espaços protegidos... ou até mesmo (quase) toda a formatação quando é selecionado o nível 8.

O Tagwipe é automaticamente utilizado na criação de projetos de tradução em ambas as ferramentas de tradução assistida por computador. Nos projetos SDL Studio, a limpeza de etiquetas de formatação é feita antes da conversão dos ficheiros originais em formato SDLXLIFF.

Nos projetos DGT-OmegaT, o Tagwipe foi integrado no DGT-OmegaT Wizard, sendo desse modo limpos todos os documentos originais em formato DOCX quando os projetos são criados ou atualizados.

No exemplo infra, o Tagwipe eliminou cerca de 90 % das etiquetas de formatação (de 121 para 14) e a segmentação foi melhorada. As etiquetas de formatação que ficaram são legítimas.

Visualização no Editor OmegaT		
	sem Tagwipe	**com Tagwipe**
Other challenges and priorities largely overlap with the achievements. Stakeholders refer to declining levels of **political commitment**, with emerging priorities, such as the refugee crisis, shifting Member States' attention away from Roma inclusion. As regards reporting by Member States, they call for more transparency, further development and closer involvement of civil society in **monitoring**, which should also facilitate peer learning. They call for more attention to **fighting discrimination**, e.g. by launching more infringement proceedings and imposing sanctions on non-compliant Member States. Regret is expressed at the fact that the Commission's Roma-targeted and mainstream policy guidance under the EU framework and Europe 2020 is **not enforceable**. There are calls to target Roma more explicitly under European and national programmes, such as the youth guarantee and Erasmus+. Many point to possible further improvements in the use of the **ESIFs** and call for NGOs and local authorities to be given direct access to funds, for better enforcement of *ex ante* conditionalities, sanctions for failure to uphold the partnership principle, better monitoring (through an increased role for the Commission and Roma themselves) and action to prevent the ineffective use of funds (e.g. training programmes not leading to employment) or their misuse (e.g. ESIF interventions financing segregated settings), including through a transparent complaint mechanism.	O<t0/>ther <t1/>challenges<t2/> and priorities <t3/>largel<t4/>y overlap with the achievements.<t5/> Stakeholders refer to declining levels of <t6/>political commitment<t7/>, with <t8/>em<t9/>e<t10/>rging<t11/> priorities, such as the refugee crisis, shifting Member States<t12/>' attention<t13/> away<t14/> from Roma inclusion. <t15/>As regards<t16/> <t17/>reporting by Member States<t18/>,<t19/> th<t20/>ey call for more transparency, further development and<t21/> closer involvement of civil society <t22/>in <t23/>monitoring<t24/>, which should also <t25/>facilitate<t26/> peer learning.<t27/> <t28/>They call for more attention to <t29/>fighting discrimination<t30/>, <t31/>e.g. <t32/>by launching <t33/>more<t34/> <t35/>infringement proceedings<t36/> and<t37/> imposing sanctions <t38/>on<t39/> <t40/>non-compl<t41/>iant<t42/> Member States. <t43/>Regret is expressed at the fact that <t44/>the Commission's <t45/>Roma<t46/>targeted<t47/> <t48/>and mainstream policy guidance under the EU <t49/>f<t50/>ramework <t51/>and <t52/>Europe<t53/> <t54/>2020 <t55/>is <t56/>not<t57/> <t58/>enforceab<t59/>le<t60/>.<t61/> <t62/>There are <t63/>calls <t64/>to<t65/> <t66/>target Roma more explicitly under European and national program<t67/>me<t68/>s, such as the <t69/>y<t70/>outh <t71/>g<t72/>uarantee <t73/>and<t74/> <t75/>Erasmus+. <t76/>Many <t77/>point to<t78/> <t79/>possible further improvements <t80/>in<t81/> <t82/>the<t83/> use of<t84/> <t85/>the <t86/>E<t87/>SIF<t88/>s<t89/> and call for NGOs and local authorities<t90/> <t91/>to be given<t92/> direct access to funds<t93/>, <t94/>for <t95/>better enforcement of <t96/>ex ante <t97/>conditionalities, sanctions for <t98/>failure to uphold <t99/>the partnership principle, better monitoring (through<t100/> an<t101/> increased role <t102/>for <t103/>the Commission and Roma themselves) <t104/>and <t105/>action to <t106/>prevent<t107/> the ineffective use <t108/>of funds <t109/>(e.g.<t110/> <t111/>training programmes not leading to employment) <t112/>or<t113/> <t114/>their <t115/>misuse <t116/>(e.g. <t117/> <t118/>ESIF interventions financing segregated settings)<t119/>, including through<t120/> a transparent complaint mechanism<t121/>.	Other challenges and priorities largely overlap with the achievements. Stakeholders refer to declining levels of <t0/>political commitment<t1/>, with emerging priorities, such as the refugee crisis, shifting Member States' attention away from Roma inclusion. As regards reporting by Member States, they call for more transparency, further development and closer involvement of civil society in <t2/>monitoring<t3/>, which should also facilitate peer learning. They call for more attention to <t4/>fighting discrimination<t5/>, e.g. by launching more infringement proceedings and imposing sanctions on non-compliant Member States. Regret is expressed at the fact that the Commission's Roma<t6/>targeted and mainstream policy guidance under the EU framework and Europe 2020 is <t7/>not<t8/> <t9/>enforceable<t10/>. There are calls to target Roma more explicitly under European and national programmes, such as the youth guarantee and Erasmus+. Many point to possible further improvements in the use of the <t11/> ESIFs<t12/> and call for NGOs and local authorities to be given direct access to funds, for better enforcement of <t13/>ex ante <t14/>conditionalities, sanctions for failure to uphold the partnership principle, better monitoring (through an increased role for the Commission and Roma themselves) and action to prevent the ineffective use of funds (e.g. training programmes not leading to employment) or their misuse (e.g. ESIF interventions financing segregated settings), including through a transparent complaint mechanism.

5. *OmegaT na DGT*

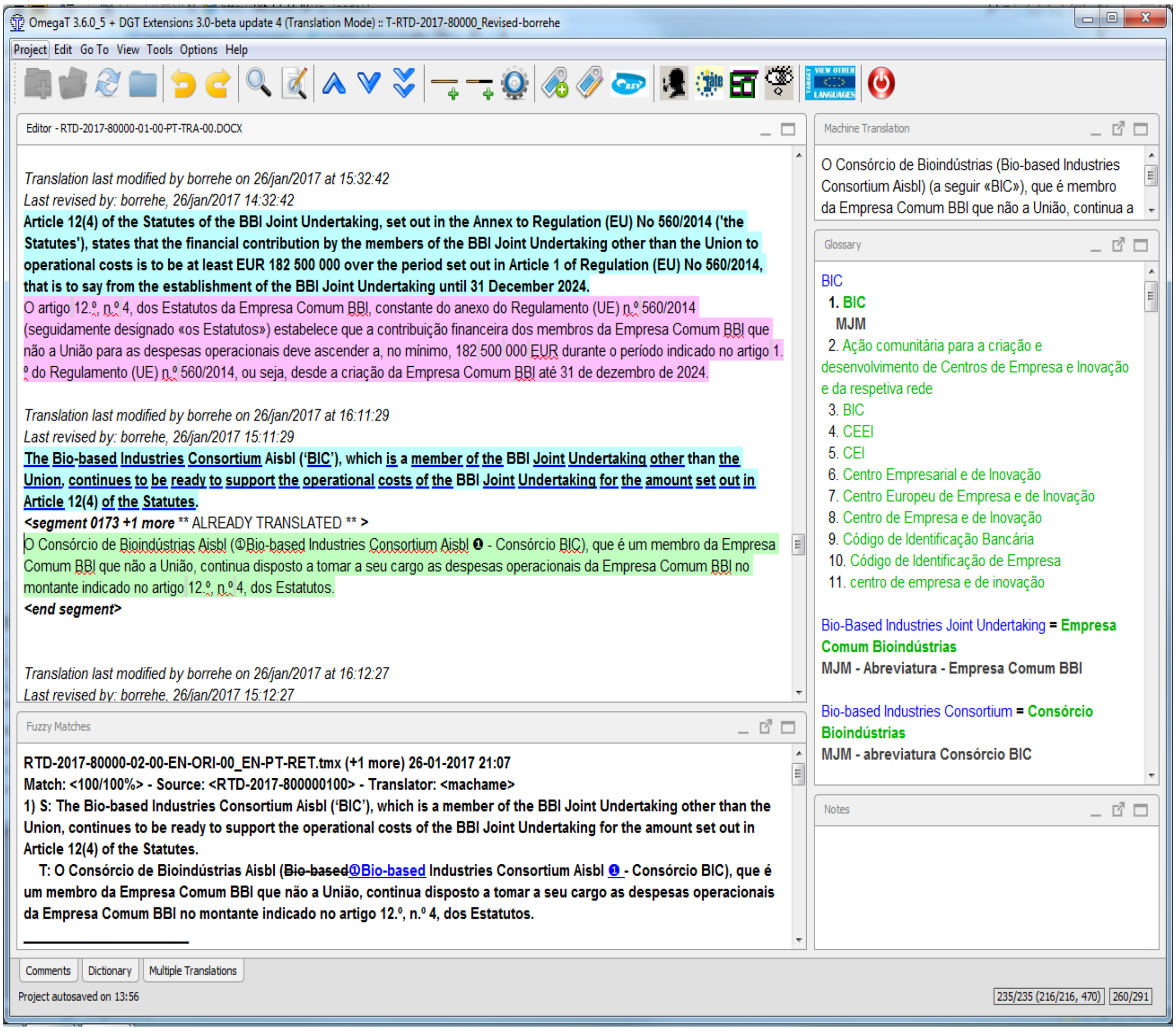
O OmegaT(10) é uma ferramenta gratuita de tradução assistida por computador desenvolvida, por iniciativa própria, inicialmente por Keith Godfrey em 2000 e que foi entretanto largamente melhorada com muitos contributos. Didier Briel é o atual gestor do projeto. Atualmente, o OmegaT é a ferramenta de tradução assistida por computador de código aberto mais utilizada.

O OmegaT na DGT é uma bifurcação da versão 3.1.2 do OmegaT com a retroinclusão de algumas funções de versões posteriores. Está também disponível uma nova versão com base no «**OmegaT standard**» 3.6.0, mas numa versão de desenvolvimento beta. Ambas as versões podem ser descarregadas aqui:

- ❑ http://185.13.37.79/ para descarregar a versão binária (fácil de instalar) e/ou o código-fonte e respetiva documentação.
- ❑ http://185.13.37.79:8003/ para programadores e comunicação de erros.

Apesar das diferenças existentes entre o DGT-OmegaT e o OmegaT, tanto quanto podemos ver, os projetos criados no DGT-OmegaT podem ser utilizados no OmegaT e vice-versa — o DGT-OmegaT mantém a interoperabilidade —, limitando-se ambos a ignorar pastas/ficheiros/informação que não «reconhecem», embora algumas funções não estejam disponíveis, como é evidente.

No sítio Web do DGT-OmegaT são apresentadas informações técnicas pormenorizadas sobre cada função nova ou adaptada. Neste artigo, limitamo-nos a resumir/destacar 10 funções principais, do ponto de vista do tradutor, que poderão ser de interesse para tradutores fora da DGT.



5.1 Barra de ferramentas da DGT

Foi internamente desenvolvido um módulo de extensão para o DGT-OmegaT que produz uma barra de ícones — equivalente a menus — que dá acesso rápido às funções mais frequentemente utilizadas, incluindo algumas que são específicas do DGT-OmegaT, nomeadamente: aplicações/bases de dados da DGT (**DocFinder**, **Quest**, **Euramis** e **IATE**), modo de revisão (**Revision Mode**) e consulta de outras línguas de chegada (**View Other Target Languages**).



Na versão em *software* livre e de código aberto do DGT-OmegaT, os ícones **DocFinder**, **Euramis** e **Quest** estão a cinzento uma vez que estas bases não podem ser acedidas de fora da DGT. A ligação para a **IATE** está ativa e dá acesso à versão pública da base[(11)].

Embora a base **Euramis** não possa ser acedida de fora das instituições da UE, uma grande parte da legislação da UE foi disponibilizada gratuitamente na Internet em formato alinhado[(12)].

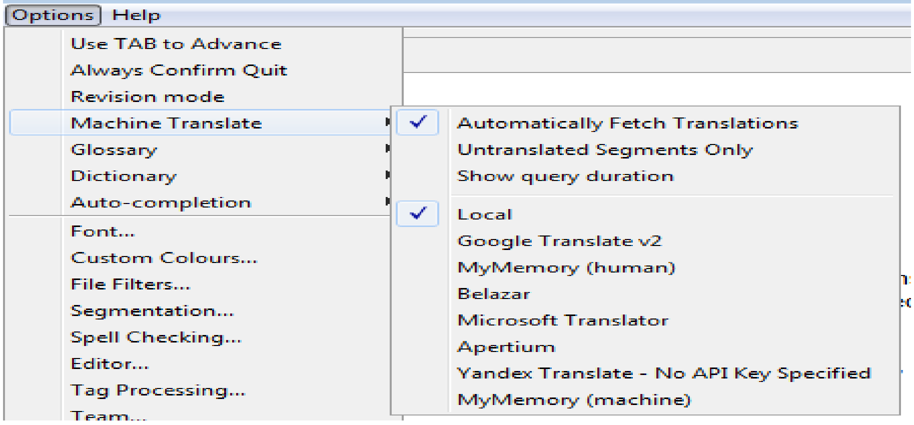
5.2 Ver ficheiros de partida e de chegada

Ao traduzir um documento, o ficheiro original do documento ativo no **Editor** pode ser aberto diretamente a partir do DGT-OmegaT selecionando **View source file** no menu **Project**. Desta forma é aberta a aplicação nativa associada ao tipo de ficheiro do documento original, que abre o documento original.

No OmegaT não há pré-visualização (em tempo real) do documento traduzido, mas o ficheiro total ou parcialmente traduzido que está ativo no **Editor** pode ser aberto a partir do DGT-OmegaT selecionando **View target file** no menu **Project**. Desta forma é aberta a aplicação nativa associada ao tipo de ficheiro do documento original, que abre o documento traduzido. Mas não há interação entre o DGT-OmegaT e a aplicação nativa, pelo que não devem ser introduzidas alterações no ficheiro de chegada na aplicação nativa uma vez que essas alterações não serão repercutidas no OmegaT!

5.3 Tradução automática

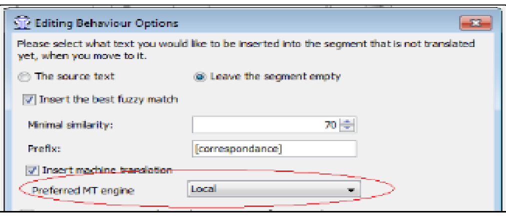
A DGT não utiliza sistemas públicos de tradução automática por diversas razões, nomeadamente por questões de confidencialidade e de direitos de autor. Em consequência, foi desenvolvida uma opção **Local** no DGT-OmegaT, que acrescenta uma nova implementação da tradução automática. Em vez de chamar um servidor em tempo real, esta implementação **Local** lê dados a partir de um ou mais ficheiros TMX. Na DGT, só o motor de tradução automática **Local** está disponível.



Na versão em *software* livre e de código aberto do DGT-OmegaT, estão disponíveis, num módulo de extensão, todos os motores da versão OmegaT 3.6, que são apresentados no menu apenas se esse módulo estiver instalado. Na versão executável publicada, este módulo está instalado.

Quando é criada uma nova subpasta **\mt** no projeto para a qual são copiados os ficheiros de tradução automática — algo que é feito automaticamente pelo DGT-OmegaT Wizard quando da criação/atualização de um projeto —, a proposta de tradução automática é apresentada no painel **Machine Translation**, em lugar de no painel **Fuzzy Matches**.

Com ou sem o módulo de extensão, a janela **Editor Behaviour Options** permite agora a inserção automática da proposta da tradução automática se não houver uma correspondência superior ao limiar de semelhança mínima definido para segmentos na memória do projeto ou nas memórias externas.



Se estiver disponível mais de um motor (ou seja, se estiver instalado um módulo de extensão tradução automática), a janela dá também a possibilidade de escolher qual o motor que será utilizado para a inserção automática.

5.4 Cores e identificação do texto sugerido no segmento aberto no Editor

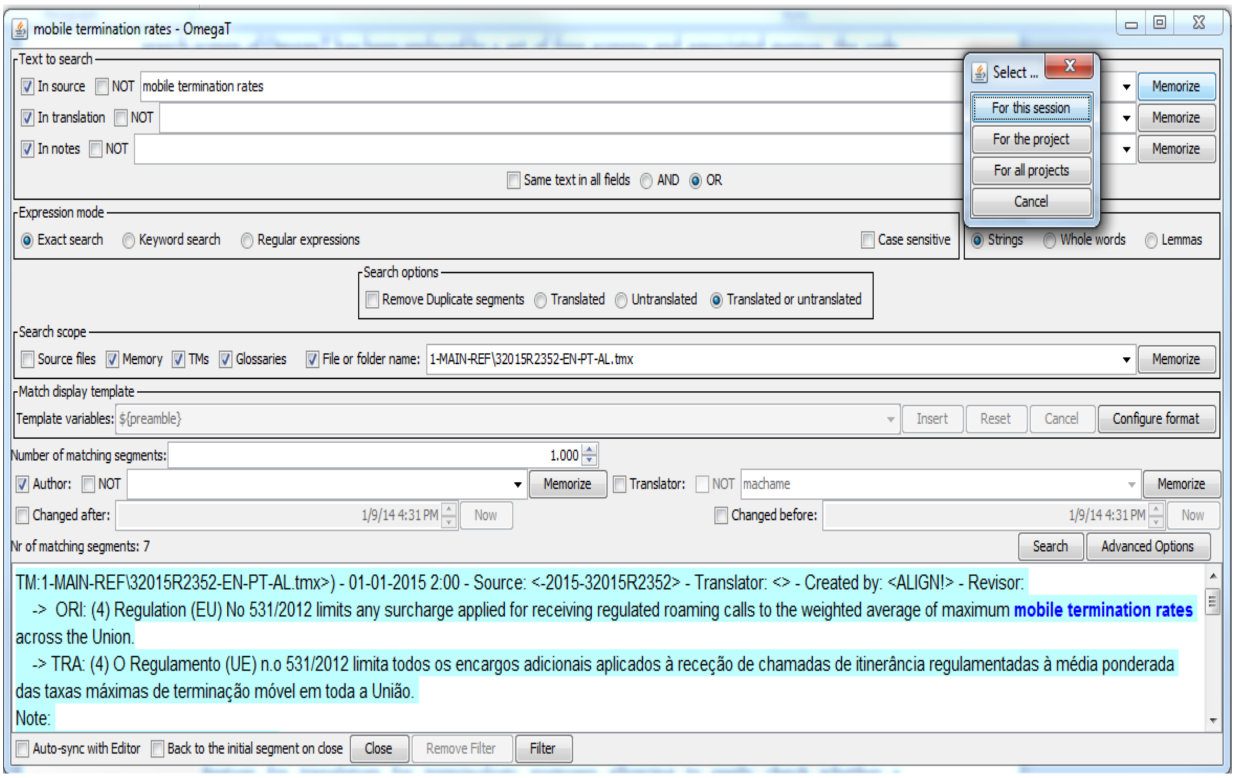
Como o OmegaT implementa uma série de mecanismos de inserção automática, quando é aberto um novo segmento é muito importante saber se a tradução proposta na zona de edição já se encontrava na memória do projeto ou se foi automaticamente inserida, mas não está validada. O OmegaT só acrescenta uma cor de fundo se o segmento provier de ficheiros TMX das subpastas **tm/auto** ou **tm/mt**. O DGT-OmegaT insere cores em alguns outros casos e também mais informações no delimitador do segmento.

Segmento vazio	**Segmento traduzido**	**Segmento pré-traduzido**
36 <segment 0440 +1 more ** EMPTY ** > <end segment>	*Translation last modified by cordoth on 16-mai-2017* Brussels, <segment 0001 ** ALREADY TRANSLATED ** > Bruxelas. <end segment>	*Last modified by unknown* All of these challenges require hug <segment 0058 ** TM/AUTO ** > Todos estes desafios exigem enormes <end segment>
Correspondência total (100 %)	**Correspondência parcial**	**Tradução automática**
GENERAL PROVISIONS <segment 0426Match 100/100/100%> DISPOSIÇÕES GERAIS <end segment>	Furthermore, the Commission will: <segment 0063 +4 more Match 100/75/66% > Além disso, a Comissão: <end segment>	See Communication 2014/C 188/02. <segment 0179 MT > Ver Comunicação 2014/C 188/02. <end segment>

5.5 Janelas de pesquisa

As funções de pesquisa do OmegaT já são muito sofisticadas. No entanto, no DGT-OmegaT, a janela de pesquisa única do OmegaT foi substituída por um conjunto de quatro janelas (Pesquisa no Projeto, Pesquisa no Diretório, Pesquisa e Substituição, Pesquisa e Pré-Tradução) e respetivos menus que continuam a ter um aspeto semelhante à do OmegaT, embora o código tenha sido completamente reescrito.

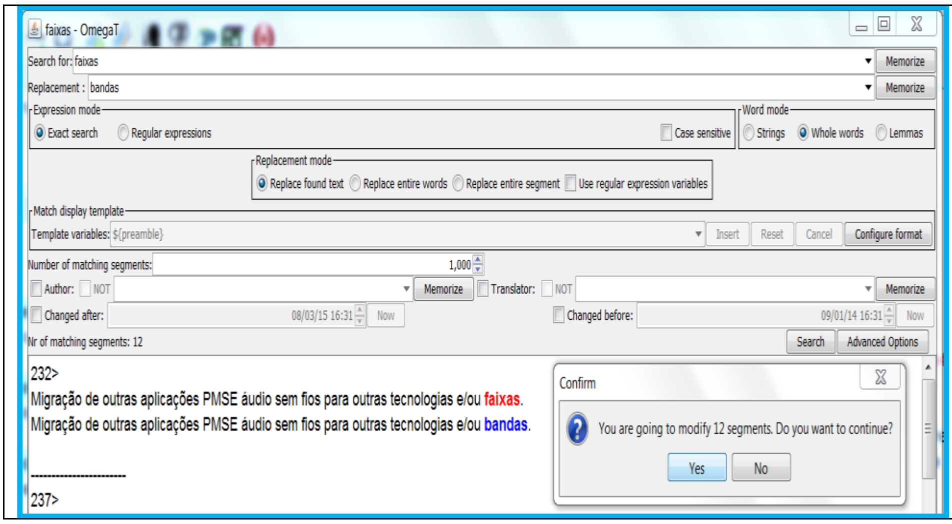
Graças a esta divisão, as janelas estão menos saturadas... e puderam ser acrescentadas mais opções. Vale realmente a pena explorar estas funções!



Algumas das novas opções de pesquisa na janela **Search Project** são:

- **Expression mode** e **Word mode**; combinados, estes modos oferecem um leque mais vasto de opções e de resultados.
- **Word mode**: As opções de pesquisa no OmegaT são baseadas em cadeias de carateres: quando se pesquisa «super», se o segmento contém «superficial», será encontrado. O DGT-OmegaT acrescenta duas alternativas:
  - **Whole words**: No modo de palavras completas, tal como em muitas outras ferramentas de edição, a amostra anterior será rejeitada a não ser que a pesquisa aceite explicitamente carateres antes e depois utilizando o caráter universal «?» (um caráter) ou «*» (0 ou mais carateres).
  - **Lemmas**: No modo lemas, são utilizados toquenizadores exatamente como para o painel de **Fuzzy Matches**: as flexões gramaticais de uma palavra serão reconhecidas, enquanto as palavras «que contenham» outra palavra não serão reconhecidas. No entanto, continua a ser uma pesquisa parcial (ou seja, uma pesquisa de um segmento que contenha determinadas palavras) sem cálculo de uma percentagem de correspondência, enquanto para o painel **Fuzzy Matches** é feita uma pesquisa completa e é calculada uma percentagem.
- **Search in** é um modo de pesquisa que foi alargado a fim de permitir uma utilização fácil dos **operadores boolianos E, OU** e/ou **NÃO**, uma função muito útil para os tradutores para fins terminológicos. Permite, em especial, verificar facilmente se um termo/expressão foi traduzido de forma coerente no projeto e/ou nas memórias externas.
- **Search by file or folder name** permite limitar a pesquisa a um ficheiro de memória de tradução (TMX) ou a uma subpasta com vários ficheiros TMX na pasta **\tm**.
- **Memorize**: Ao contrário do que acontece no OmegaT, a memorização de uma pesquisa é feita através de um botão (para permitir uma memorização seletiva apenas de pesquisas relevantes) e há opções para memorizar a pesquisa para todos os projetos, para o projeto que está aberto ou apenas para a sessão em curso. Além disso, encontram-se já memorizadas algumas expressões regulares frequentemente utilizadas.
- **Match display template**: Com o modelo de visualização de correspondências, é possível configurar a forma de visualização (como para as correspondências) utilizando a opção **Config format**.

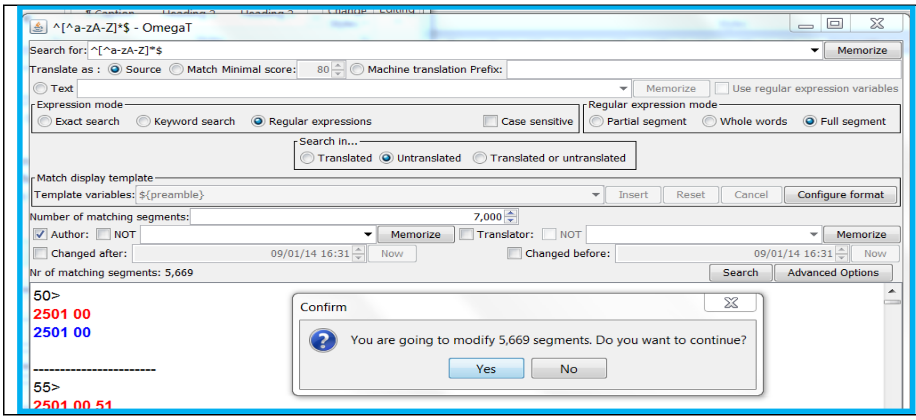
A janela **Search and Replace** segue a mesma abordagem da janela **Search Project**.



Foi adaptada de modo a incluir um painel de resultados, permitindo assim verificar os segmentos afetados antes de lançar a operação de substituição propriamente dita. Isto é muito útil uma vez que não existe a opção de anular a operação (**Undo**)!

A janela **Search and Pre-translate** é específica do DGT-OmegaT. A sua função é preencher rapidamente segmentos cuja tradução é irrelevante (por exemplo, números e outros carateres não traduzíveis) ou mais facilmente executada numa operação que abrange vários segmentos. Funciona de forma semelhante a **Search and Replace**, mas executa a pesquisa no **segmento de partida**. Por conseguinte, pode ser usada para segmentos não traduzidos.

Enquanto a janela **Search and Replace** apenas permite traduzir com texto fixo, a pré-tradução permite preencher segmentos com tradução proveniente de: **Source**, copiando o segmento de partida para o segmento de chegada; **Match**, pré-traduzindo a partir de correspondências de memórias externas (o limiar de semelhança pode ser definido) e **Text**, tal como para **Search and Replace**, permite a utilização de uma cadeia de carateres constante com a possibilidade de usar variáveis reconhecidas durante a pesquisa.

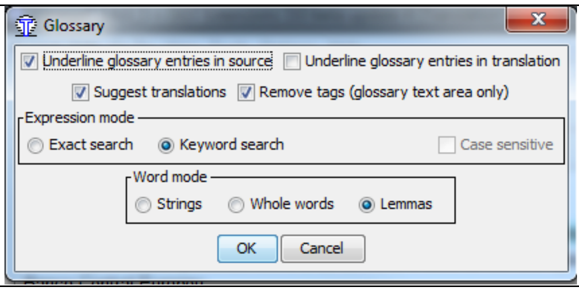


Este é um exemplo da sua utilização com **Regular Expressions** para pré-traduzir segmentos sem carateres alfabéticos (principalmente números).

A janela **Search Directory** segue uma abordagem similar, com adaptações aos seus fins específicos.

<u>5.6 Glossário</u>

O painel **Glossary** utiliza cores para diferenciar o termo de partida, o termo de chegada e o terceiro campo (ver captação de ecrã no ponto 4), sendo o último campo apresentado juntamente com o respetivo termo de chegada. Além disso, as entradas após as entradas do glossário de escrita são apresentadas por ordem alfabética, tornando mais fácil a consulta.



O submenu **Glossary — Glossary Pane and Transtips configuration** oferece ao tradutor algumas opções importantes para a visualização da terminologia e sincroniza os algoritmos de pesquisa de palavras a fim de produzir resultados coerentes em ambos os painéis.

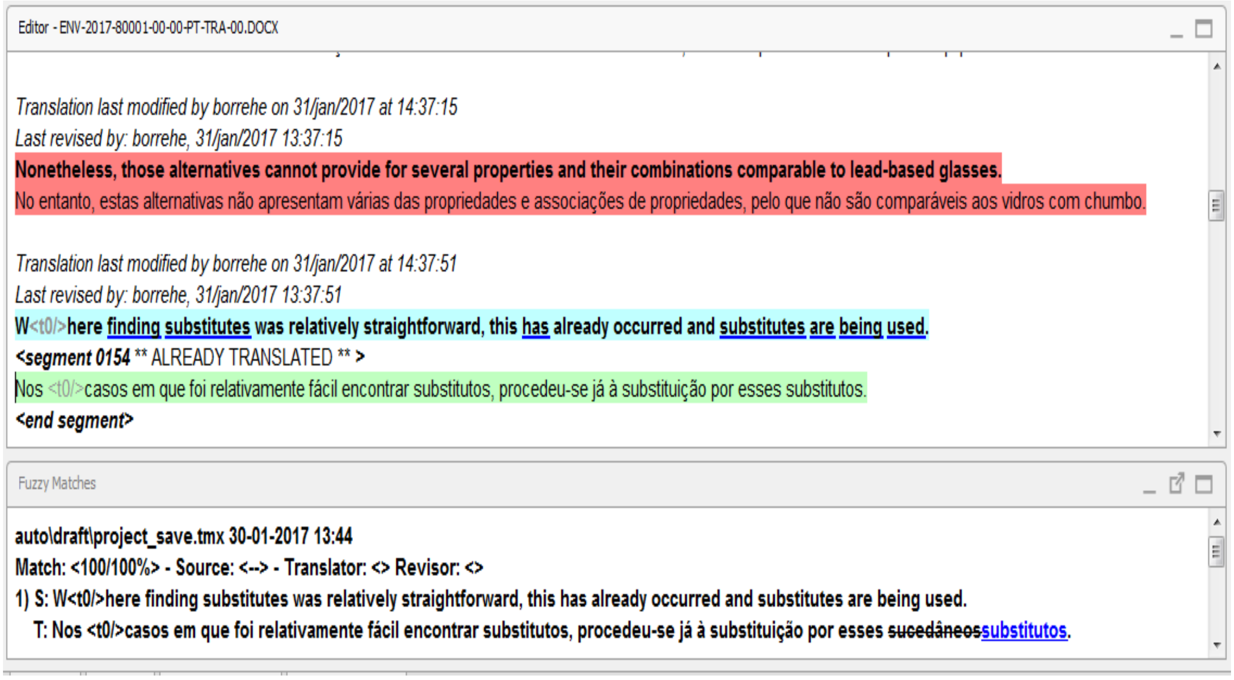
<u>5.7 Cálculo de estatísticas de correspondências</u>

O cálculo das estatísticas de correspondências — **Match Statistics** e **Match Statistics per File** — são efetuadas em operação de fundo, permitindo assim ao tradutor começar a traduzir enquanto as estatísticas estão a ser geradas. Isto é especialmente interessante em grandes projetos de centenas/milhares de páginas com muitas memórias externas, uma vez que o cálculo de estatísticas de correspondências (por ficheiro) nesses projetos pode demorar bastante tempo.

<u>5.8 Modo de revisão</u>

Na DGT, os tradutores podem ser também revisores e, de um modo geral, o tradutor tem a última palavra, o que significa que, depois de um documento ser revisto por um colega tradutor na qualidade de revisor, o tradutor verificará as alterações efetuadas e pode aceitar ou rejeitar as mesmas. Tendo isto em conta, foi desenvolvido um fluxo de trabalho de revisão no DGT-OmegaT Wizard e foram desenvolvidas as seguintes funções no DGT-OmegaT.

No DGT-OmegaT, há agora dois modos — **modo de revisão** e **modo de tradução** — e os segmentos têm três estados: não traduzido, traduzido e revisto.



Quando o revisor seleciona o **Revision Mode** :

- Os segmentos abertos e de qualquer forma validados no **Editor** — alterados ou não — são marcados como **Last revised by**, com o nome do revisor. Esta informação é apresentada para esses segmentos no **Editor** e, se estiverem assim configuradas, nas janelas **Fuzzy Matches** e **Search**. Esta marca não pode ser eliminada... mas pode ser efetuada uma exportação sem essa marca.
- Se se copiar a memória do projeto não revisto para a pasta **\tm** (identificada como «sem revisão», por exemplo), as marcas de revisão são apresentadas no segmento traduzido desde que a nova opção **View diff in target** esteja ativada nas opções do menu **External TMX**.
- Há também uma nova opção **Go to Next Unrevised Segment** com o atalho **Ctrl+U** (o mesmo utilizado no **Translation Mode** para **Go to Next Untranslated Segment**), que permite abrir apenas os segmentos ainda não revistos.

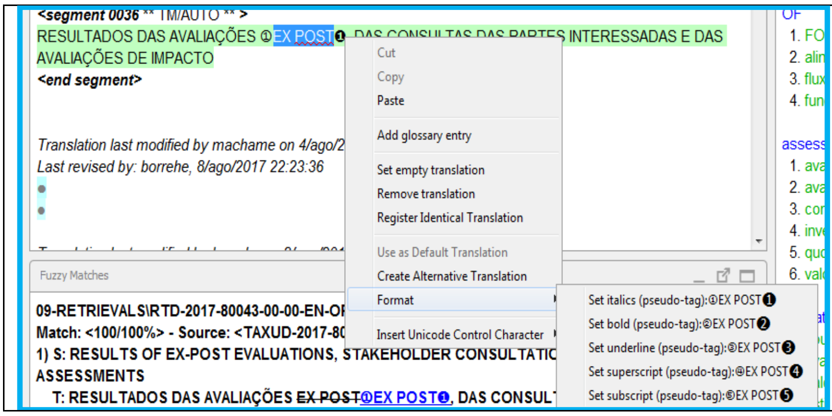
Quando recebe o projeto revisto para finalização, o tradutor pode fácil e rapidamente verificar os segmentos alterados pelo revisor com as seguintes duas novas opções:

- **Mark Revised and Changed Segments** no menu **View**, sendo os segmentos alterados pelo revisor mostrados com um fundo vermelho.
- **Go To Next/Previous Revised & Changed Segment** (e respetivos atalhos), abrindo assim sucessivamente para verificação apenas os segmentos alterados pelo revisor.

O modo de revisão pode também ser utilizado pelo tradutor quando determinado projeto não é revisto por um colega, sendo nesse caso o tradutor o seu próprio revisor, como acontece num número substancial de documentos.

5.9 Pseudoetiquetas de formatação

No DGT-OmegaT, há também uma opção para acrescentar etiquetas de formatação no segmento de chegada que não existe no segmento de partida, utilizando a função **Format** (**pseudo-tags**).



Selecionando o texto a formatar, clicando no botão direito do rato e escolhendo o formato desejado (itálico, negrito, sublinhado, sobrescrito e subscrito), são inseridos carateres existente na tabela Unicode mas que são raramente utilizados.

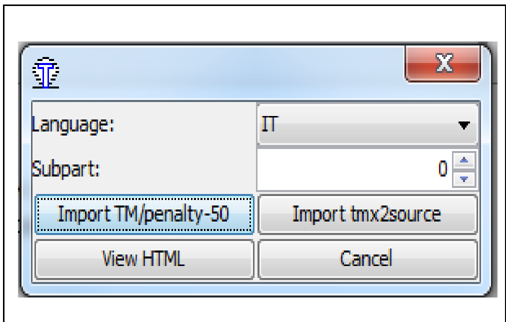
Quando os documentos traduzidos são gerados, o *script* **Reformatter** transforma essas pseudoetiquetas no formato correspondente no documento de chegada. No entanto, as pseudoetiquetas devem ser utilizadas apenas quando forem indispensáveis, visto que, por vezes, pode ser afetada outra formatação do segmento. Não é a solução perfeita (e, de momento, apenas funciona para ficheiros DOCX), mas pode ser muito útil!

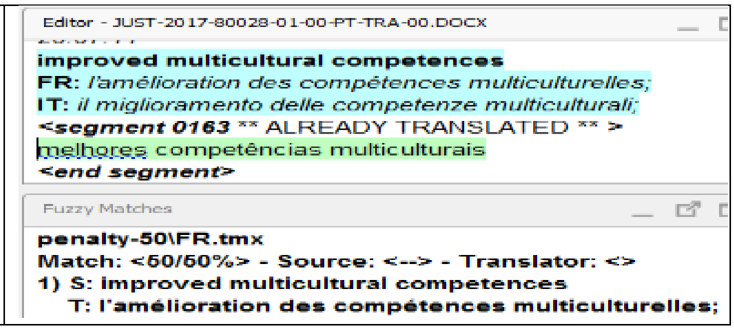
5.10 Consulta de outras línguas de chegada (concordância interlinguística)

Muitos dos documentos da UE são traduzidos simultaneamente para algumas/todas as línguas oficiais da UE. Para fins de coerência multilingue, é útil que os tradutores possam olhar para as traduções em curso nas outras línguas. O OmegaT dispõe já de um boa funcionalidade para esse efeito: a pasta **tmx2source**, mas o problema é que os ficheiros têm de ser pedidos e copiados manualmente para essa pasta.

Para esse efeito, foi feita a seguinte adaptação na instalação do SDL Studio na DGT: sempre que uma tradução (em curso, não final!) é guardada, o ficheiro SDLXLIFF é copiado para uma pasta comum. Em qualquer momento, o tradutor desse mesmo documento para uma outra língua pode utilizar um menu contextual (módulo de extensão no Studio) que irá converter o ficheiro SDLXLIFF em HTML e visualizá-lo no navegador.

No DGT-OmegaT, **View Other Target Languages** —  — combina esta com outras opções. Em primeiro lugar, funciona nos dois sentidos — os utilizadores do DGT-OmegaT podem ver os documentos em tradução em Studio e vice-versa. Em segundo lugar, os utilizadores do DGT-OmegaT podem escolher que esses ficheiros sejam automaticamente convertidos para TMX (não para HTLM) e copiados para a pasta **tmx2source** (para visualização no **Editor**) e/ou para a pasta **tm/penalty-50** (para visualização no painel de **Fuzzy Matches**).





N.B.: Neste artigo são utilizados os termos em inglês das funções e menus do DGT-OmegaT uma vez que a versão utilizada na DGT é a versão inglesa e neste artigo utilizamos captações de ecrã dessa versão. No entanto, o OmegaT está disponível em 32 línguas, entre as quais o português (de Portugal e do Brasil).

(Continua na parte 2...)

Maria.Machado@ec.europa.eu
Thomas.Cordonnier@ext.ec.europa.eu
Hilario.Fontes@ec.europa.eu

---

(1) «*Software* livre e de código aberto — um bom amigo do tradutor», *in* «a folha», n.º 45 — verão de 2014, http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_pt.pdf.
(2) OmegaT, http://omegat.org/pt/.
(3) OmegaT, *Descarregamentos*, http://omegat.org/download.
(4) Comissão Europeia, *Research & Innovation: Research and Innovation funding 2014-2020*, http://ec.europa.eu/research/fp7/index_en.cfm.
(5) Comissão Europeia, *Horizon 2020*, https://ec.europa.eu/programmes/horizon2020/.
(6) *European Commission Open Source Strategy*, http://ec.europa.eu/dgs/informatics/oss_tech/index_en.htm.
(7) Direção Geral da Tradução, *Translation tools and workflow*, 2017, https://publications.europa.eu/en/publication-detail/-/publication/00e51a8e-9c50-11e6-868c-01aa75ed71a1.
(8) Comissão Europeia, *Machine translation for public administrations — MT@EC*, https://ec.europa.eu/info/resources-partners/machine-translation-public-administrations-mtec_en (disponível apenas para as administrações públicas).
(9) Moses, http://www.statmt.org/moses/.
(10) Wikipédia, *OmegaT*, https://pt.wikipedia.org/wiki/OmegaT.
(11) União Europeia, *IATE*, http://iate.europa.eu/about_IATE.html.
(12) Fontes, H. L., «*Corpora* multilingues da União Europeia para reutilizar na tradução», *in* «a folha», n.º 45 — verão de 2014, http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_pt.pdf.

Tendências da língua portuguesa: as inócuas e as iníquas (VI)

Jorge Madeira Mendes
Direção-Geral da Tradução — Comissão Europeia

As ciências naturais têm vertentes intrigantes.

Recentemente, foi-me dado saber que, em matéria de antílopes, há os antílopes *fêmea* — pelo que é de presumir haver também os antílopes *macho*.

E as enguias começam por ser as *bebé*. Suponho pois que, à medida que forem avançando em idade, passarão sucessivamente a enguias juvenil, enguias adolescente, enguias adulto, enguias madurito, enguias serôdio, enguias vetusto... por aí fora, até chegarem a enguias velho, enguias moribundo, enfim, enguias falecido.

Depois, temos os raios *ultravioleta*, o que me leva a supor termos também raios *infravermelho*.

Aventurando-nos agora pelo domínio da sociologia, eis que nos aparecem os homens *gay*. E os «outros»? Serão os homens *heterossexual*?

Estes frustrantes exemplos, extraio-os dos canais televisivos ditos «temáticos», onde as legendagens e dobragens resultam, com frequência infelizmente excessiva, de traduções amadorísticas e apressadas.

E é, desde logo, manifesta a presença mal disfarçada do inglês (sem desprimor para esta bela língua, inocente do mau uso que dela fazem). Se o original diz *female antelopes*, o tradutor amador (para o qual «traduzir» parece querer dizer «substituir acriticamente cada palavra da língua de partida por algo que se lhe assemelhe na língua de chegada») não duvida de que há que pôr «antílopes fêmea»; pois não está o adjetivo *female* no singular? Se o original diz *baby eels*, o tradutor amador põe «enguias bebé»; pois não está o adjetivo *baby* no singular? Se o original diz *ultraviolet rays*, o tradutor amador põe «raios ultravioleta»; pois não está o adjetivo *ultraviolet* no singular? Se o original diz *gay men*, o tradutor amador põe «homens *gay*»; pois não está o adjetivo *gay* no singular?

No caso das «enguias *bebé*» e dos «antílopes *fêmea*», a catástrofe é tão mais grave quanto vem adicionar-se a uma outra tendência em alastramento: os nomes de espécies biológicas, quando compostos, formarem o plural derivando apenas o primeiro termo e mantendo invariável o segundo («o tubarão-baleia, os tubarões-*baleia*»; «a couve-flor, as couves-*flor*».

Leiamos os entendidos:
– «Nos compostos formados por dois nomes, ambos os elementos se usam no plural: *couves-flores* (...) Nos compostos formados por nome e adjetivo ou por adjetivo e nome, ambos os elementos se usam no plural: *amores-perfeitos*, *capitães-mores*, *gentis-homens*, etc.»[1].
– «Quando os termos componentes [dos substantivos compostos] se ligam por hífen, podem variar todos ou apenas um deles: couve-flor, couves-flores/ guarda-marinha, guardas-marinha...»[2].

É compreensível que, num substantivo composto como «guarda-marinha», o plural afete apenas o primeiro elemento (guarda/ guardas), porquanto o segundo não varia (a marinha é só uma).

Ora, nos exemplos que indiquei, não estamos sequer em presença de substantivos compostos, com os dois elementos ligados por hífen e em que o primeiro seria um substantivo e o segundo um adjetivo ou outro substantivo. O que se passa é haver um substantivo qualificado por um adjetivo — tão simples quanto isso. E a regra impõe que, em português, o adjetivo qualificativo concorde em género e em número com o substantivo qualificado. Os adjetivos *bebé*, *fêmea* e *gay*, embora invariáveis em género, são bem variáveis em número. De qualquer modo, faltaria que uma larva de enguia passasse agora a constituir uma espécie no seu pleno direito...

Ainda não chegámos ao ponto de ouvir (ou ler) «os cidadãos *tranquilo* apreciam cidades pouco *movimentado*; e os homens bem *apresentado* atraem as mulheres *bonito*». Mas, por este andar, lá chegaremos.

Ignoro se «pérolas» destas correspondem a uma tendência, ou se se tratará apenas de vulgares pontapés na gramática, como tantos que sempre se ouviram. Com efeito, afirmamos amiúde que «cada vez mais isto» ou «cada vez mais aquilo», quando, na verdade, se trata de situações de longa data, com a principal diferença de que, outrora, tais pontapés eram menos divulgados pelos meios de comunicação social, desprovidos da hodierna capacidade de penetração e também sujeitos a crivos redatoriais mais exigentes.

É possível. Mas, a haver aqui tendência, será, incontestavelmente, iníqua. E bastante me inquieta a facilidade e a velocidade de insinuação destes erros no discurso corrente.

***

Começo também a ler cada vez mais frases como a do seguinte exemplo: «Já lá *tivemos* várias vezes; portanto, se você *tiver* disposto, pode também ir.»

Trata-se manifestamente da erosão do verbo «estar» para «tar». No entanto, é raro ver-se, em forma escrita, «tarão lá num instante» ou «tenho tado». E a razão é discernível: as palavras «tarão» ou «tado» não existem em português; embora comuns na oralidade, estão normalmente ausentes do registo

gráfico. Em contrapartida, as palavras «tivemos» e «tiver» são reais em português, se bem que se trate de formas de um outro verbo — *ter*.

Poderá estar a desenrolar-se um fenómeno evolutivo na língua portuguesa, em que o verbo «estar» passa, por erosão, a «tar», apoiando-se a tendência na semelhança entre algumas das suas formas erodidas e as correspondentes formas de um outro verbo, esse legitimado.

Ora, do mesmo modo que, de «general», se passou, por erosão, para «geral»; do mesmo modo que, de «tener», se passou, por erosão, para «ter» (depois de uma transição por «teer»); do mesmo modo que, de «preguntar», se passou, por deturpação de pronúncia, para «perguntar» (e note-se que a forma «preguntar» só terá sido oficialmente abandonada há cerca de um século, o que, na história praticamente milenar da língua portuguesa, é recente) — a substituição de «estar» por «tar» (com a correspondente conjugação: «eu tou», ou «tô», «tu tás», «ele tá», «nós tamos», «vocês, e eles, tão», «tarei», «taremos», «taria», «taríeis»...) poderá ser, pacificamente, encarada como uma tendência inócua. Com efeito, enquanto fenómeno meramente fonético, não envolve atropelo da sintaxe.

E consta-me que, pelo menos em Moçambique, é já preconizada a avalização do novo verbo «tar».

Jorge-Madeira.Mendes@ec.europa.eu

---

(1) Bergström, M.; Reis, Neves, *Prontuário Ortográfico e Guia da Língua Portuguesa*, 50.ª edição, Casa das Letras, 2011.
(2) Cunha, C.; Cintra, L. F. L., *Breve Gramática do Português Contemporâneo*, Edições João Sá da Costa, Lisboa, 1985.

Sardinheiras e cravos

Ulrike Hub
Direção-Geral da Tradução — Comissão Europeia

Caros leitores,

É isto uma continuação, um «P.S.-zinho», ao artigo «Aventuras e desventuras de uma tradutora alemã entre portugueses» (aparecido n'«a folha» de outono de 2016). Eu, a alemã tradutora na Comissão Europeia em Bruxelas que já tinha feito um *office swap* em 2015/16, tive a oportunidade e o prazer inesperados de poder voltar a passar uns meses no departamento português.

Com pouco tempo para estudar em casa, *je me suis vite heurtée aux limites difficiles à franchir* sem aulas de português. Não obstante, durante a minha estadia fiquei a saber mais algumas coisas sobre Portugal, e o português «entrou-me» mais definitivamente na alma. Assim, quando vi uns miúdos japoneses descer do autocarro escolar com os cartazes japoneses deles, esquisitos, lindos, todos iguais, só de cores diferentes, o que pensei foi: «Gosto!!» (e não «*Oh, schön!*» ou «*Joli, ça!*»).

Tomei nota de mais algumas coisas que descobri e aprendi (ou não), para os que estejam interessados. Como sempre, em várias línguas, que só em português ainda não consigo. Visto a tradição de trazer bolos alemães, quis criar um «bolo bola» de frases — *sous forme d'une spirale, rond, avec les phrases qui tournent*. Infelizmente não consegui, ou talvez felizmente, visto que o que criei *aurait donné un torticollis*...

Durante os cinco meses do *swap*, comi doce de tomate, e bacalhau, deliciosos. Vi uns episódios do «Conta-me como foi», andei «desencontrada» com uma colega, fui tomar café com três portugueses ao

mesmo tempo e percebi o que disseram... (*mais s'ils avaient parlé juste entre eux, je n'y serais pas arrivée*). O erro da secretária já é «*history*», mas dei erros novos, assim disse «seize» em vez de «dezasseis» (*c'est bien* «quinze» *en français ET en portugais, alors pour quelle raison est-ce «seize» en français, mais pas en portugais ?!*).

Escrevi o *e-mail* mais curto da minha vida: «É!»

Descreveram-me umas torradas em três partes, com muita manteiga. Parece que todos queriam sempre a parte do meio ☺. Levei algum tempo a perceber que é num bar que se comem, hei de encomendar quando for a Portugal. Soube também dos três arrozes típicos (de polvo, de pato e de tamboril). Aprendi que os sapatos se arrumam no quarto, e que há escolas pintadas de cor de rosa.

Descobri que o dicionário da Academia das Ciências tem dois tomos, *both equally thick*, mas que estranhamente o primeiro tomo só tem as letras A-F (6 letras) e o segundo tem G-Z (20!!). É que em português devem existir muitas palavras que comecem com A, B, C, D, E ou F.

Contaram-me impressões das relações entre portugueses e espanhóis, e disseram-me que, falando português, podia fingir que falo russo (sempre que os outros presentes não falem português nem russo, claro). Talvez tente uma vez — vamos lá ver... Deram-me a ler livros — de uma autora canadiana... que já é uma das minhas preferidas, um livro alemão ☺, um «Reiseführer» alemão sobre Portugal, e li contos dos irmãos Grimm em português, o que — apesar de me parecerem mais lindos em português do que em alemão — foi «escanifobético» (*hopefully I'm using it correctly, this word*).

Gostei muito desse «escanifobético». Achava que o Mário de Carvalho tinha inventado a palavra, mas parece que não, existe mesmo. E as «sardinheiras»! Pegando no *Casos do Beco das Sardinheiras*, estava à espera de histórias com mulheres vendedoras de sardinhas. Depois achei que as «sardinheiras» talvez fossem barcos para a pesca de sardinhas. Mas barcos também não apareceram nessas histórias, maravilhosamente estranhas!... Tive de procurar no dicionário, que isso é trabalho!, e «*it turned out that*» sardinheiras são «*geraniums*»... (?!)

A outra flor que me fez estranhar foi o «cravo». Como é que uma única palavra pode ter tantos significados diferentes? *At least: carnation, nail, harpsichord, wart*. (Sugiro que utilizem o «fuz» para um destes significados?)

Achei admirável o «Já ia a ir-me embora» (*quelle précision!*), adoro o «Mau, mau!» (por acaso é também o nome de um jogo de cartas na sul da Alemanha), o «nadinha de nada», tal como o tipo de *t-shirt* que se chama «tomara que caia» (no Brasil), parece que é «caicai» em Portugal. Pode morrer-se de morte matada! E não só se renovam as casas, mas também os passaportes (eu vejo *les hommes de métier avec leurs outils qui s'attaquent aux passeports*).

Além disso, gostei das «tagarelices», da «semínima», do «aqui há gato», do «redemoinho» (*je vois un démon danser*), da «água que passarinho não bebe» e da «boa pinga», do «súper» = «supermercado», da «boa boca», e dos «alergénios» que não são «génios». Adoro o «percevejo» (a palavra)! (Vem de «percebo» e «vejo»? Imaginei: Está a falar uma pessoa, e depois de ter dito «Perce...» é picada pelo percevejo. Faz-se um silêncio pequenino, depois continua a pessoa a falar, mas dizendo «vejo» *à la place de* «percebo». *Bon, c'est chercher loin, je l'admets!!*).

O português pode ser flexível, um pouco como o inglês, que tem verbos como «*to tango*», ou o neerlandês, com «*tennissen*». Em português podem «ultimar-se os preparativos» e pode-se «sambar». Mas é verdade, em alemão temos «*duschen*» e «*grasen*».

Em português, pronunciam-se diferentemente as cidades de M**ó**naco (PT) [Monac**o** (FR), Mon**a**co (DE)], e Vars**ó**via (PT) [Varsov**i**e (FR), W**a**rschau (DE)]. Descobri que acrescentam um A a «Heidelberg**a**» ☺.

Com «bossa nova» vi na mente uma nódoa negra, recente. Quando os colegas portugueses falaram de «Viana», eu entendi «Vienna». E li «toro competente» mas era «foro competente»...

O «vespertino» fez-me lembrar a vespa que tinha quando jovem. A «espelunca» é como a «*Spelunke*» alemã. Mas a «paragem cardíaca» não é uma paragem de autocarro, «antigamente» ou «antigo» não foi na antiguidade, e «*granda* lata» não quer dizer «que sorte» (eu tinha a certeza... mas absoluta!... e traduzi mal...). Percebi que um fim de semana «comprido» é muito melhor do que um fim de semana «cumprido» (mas pronuncia-se quase da mesma maneira, se não estiver errada). E que «mais daqui a um pouco» não quer dizer que haverá mais (rebuçados, ou outra coisa boa qualquer) dentro de pouco tempo.

«Estou esquecido» não quer dizer que a pessoa que fala foi esquecida. O «sótão» português é o lugar mais alto de um prédio, mesmo que em Espanha o «*sótano*» seja o lugar mais baixo. «Enxugar» é o contrário de «enxaguar», e «puxe» é o contrário de «*push*».

Porque é «figueir**a**» mas «limoeir**o**»?

Gostei das «águas-furtadas» (os vizinhos furtaram a água, uns aos outros, lá em cima? é um país seco... ai, a minha imaginação às vezes vai longe de mais).

Demais, de mais, amais, a mais — existem todos? Ainda não percebi bem a diferença... Mas já sei que sempre tenho de *acknowledge that my kids are MY kids*: «os MEUS filhos» (e não simplesmente «os filhos», quando falo dos meus, como em francês e em alemão, «*les enfants*», «*die Kinder*»).

Percebi, logo na televisão, mais ou menos bem, a canção da Eurovisão (a dos irmãos Sobral), mas ao ouvir falar da Volta à França no telejornal da RTPI, não percebi em que lugar ficou o melhor português — os números ordinais são difíceis.

Tentei «bater as claras em castelo». Ainda vou ter de praticar ☺.

Os meus filhos ainda não falam português, mas percebem «Não sei» e «Não faço ideia». E cantam a canção da Eurovisão — no «português» deles.

Procurando no IATE, agora vejo na mente um iate no mar, com um céu azul, vento e nuvens brancas, e fico feliz.

Continuo a dizer «dentro de três dias» (em vez de «daqui a três dias») e «até à semana que vem» (em vez de «até para a semana»). E «sente bem» em vez de «cheira bem». Confundo, obstinadamente, o «o» e o «ou». E os acentos — nem quero falar do assunto. Desisto do «com ou sem artigo», nos países — «em Portugal», mas «na Itália», etc. Desisto também de produzir o «eeeeee» em «eles» — tão comprido. *And I'm afraid I will never get the* «tempos verbais» *right*.

Foi giro poder ficar com os colegas portugueses mais uma vez. Teria gostado de preparar um *Eiskaffee* para os que quisessem, é um café frio com gelo, gelado e natas — coisa típica dos verões quentes (relativamente... 30-35 graus) do sul da Alemanha, mas quando esteve calor eu estive de férias (sorte minha ☺).

Li, já não sei onde, que «atrás de tempo, tempo virá». Uma colega já me disse várias vezes que «a vida dá muitas voltas». E não foi o Saramago que disse que o destino não é uma linha reta? Talvez ainda possa voltar, um dia, para mais uns meses. Fica o *Eiskaffee* para então.

Beijinhos grandinhos!

Ulrike.Hub@ec.europa.eu

Desportuguesmente

Philippe Magnan Gariso
Tradutor técnico — Mota-Engil, Engenharia e Construção, SA

Pensei, pensei, pensei, voltei a pensar, dei voltas à cabeça, quase bati com a cabeça nas paredes, e zás, não é que do infortúnio se fez luz! Tão intensas que eram as chamas, tão más as reportagens que davam conta do seu avanço pelas serras, que dei comigo a comentar: vale mesmo a pena esclarecer alguns conceitos, corrigir algumas formas de dizer, na esperança de que alguém oiça ou, melhor, leia, nessas longínquas escolas de jornalismo e nesses recônditos estúdios em que se lêem telejornais e jornais radiofónicos.

Pois bem, o leitor já percebeu o que me traz à conversa: fogos, incêndios, frentes, operacionais, chamas controladas, dominadas, em fase de resolução e resolvidas, as fitas do tempo e, ainda, as vítimas mortais. Comecemos então pelo princípio. Não faltaram espaços noticiosos sobre a malfadada época dos fogos florestais em Portugal. Em primeiro lugar, não há «fogos» florestais, mas sim «incêndios» florestais: E porquê? Recordando o que aprendi no serviço militar, na protecção civil, em França, um fogo pode ser extinto com um copo de água ao fim de um minuto; ao fim de cinco minutos é preciso um balde de água; para lá desse tempo, se o fogo não tiver sido extinto, ou seja, se estiver descontrolado, passa a designar-se por «incêndio» exigindo outros meios e procedimentos de combate. Há especialistas que apresentam outros tempos, por exemplo, ao fim de três minutos o fogo combate-se com um balde de água, ao fim de cinco já é necessário uma cisterna, e a partir daí o fogo passa a incêndio, se não tiver sido extinto.

Ora, o incêndio vai lavrando e desenvolve-se em frentes a que as gentes do jornalismo teimam em qualificar de «activas»; as «frentes activas». Mas, se a frente arde é porque, forçosamente, está «activa», senão não haveria frente. Pois é, amigo leitor, há frentes de incêndio, e se as há é porque estão «activas»; daí estarmos perante uma redundância, tal como a frase, por exemplo: «Ainda há dois, três, quatro, ou seja quantos forem, incêndios activos no norte do país». Então, se não estivessem activos, não haveria incêndios, não será? Há um «x» números de incêndios; e se os há estão activos. Parece-me! Não concordará o leitor?

Nessas reportagens, ouvimos dizer que os operacionais estão apoiados e suportados por «x» meios terrestres. Apoiados e suportados? Qual será a diferença? Confesso-me incapaz de destrinçar. Atiçando mais as labaredas, dizem-nos que o fogo está controlado» ou «dominado». Não é o fogo que está dominado, ou controlado; na realidade, o que acontece é que numa dada fase do combate ao incêndio se conseguiu delimitar, ou confinar, o perímetro que naquele instante arde. Para espanto dos especialistas, essas gentes da comunicação dita «social» emprega, pelo menos este ano, uma expressão nunca antes ouvida no meio: o fogo está em fase de «resolução». Concretamente, o que significa isto? Como se caracteriza um fogo, melhor dizendo um incêndio, em fase de resolução? Desde Junho, com Pedrógão Grande, que não se ouviu, uma única vez que fosse, que um incêndio estivesse «extinto». Porque será? Apesar de muitos desses incêndios terem sido, de facto, extintos. E o que pensar de um incêndio resolvido? Este «resolvido» inclui o rescaldo, ou não?

Já agora, o que pensar, das «fitas do tempo», a propósito das falhas nas comunicações. «Fita» porquê? Em bom português, sempre se falou em linha do tempo, cronologia, histórico dos acontecimentos… não faltam soluções. Em inglês, é corrente ouvirmos e lermos a expressão «*time line*», pelo que não vale a pena inventar, porque a tradução literal, neste caso, serve perfeitamente. Finalmente, as «vítimas mortais». Do ponto de vista morfológico, «vítima» é um substantivo; perdão, um nome, como agora se diz; «mortal» pode ser um nome (o mortal, aquele que morre, ou um adjectivo, ou seja, o que causa morte, que mata, que é letal. Então, em «vítima mortal» estamos a falar de uma vítima que causa morte, ou de um sujeito que morreu, que está morto? Se considerarmos «mortal» como adjectivo, quão más são as vítimas que são mortais! Matam, causam morte! Se, ao invés, considerarmos como um nome, então estamos perante um vítima que morre. Li, algures, que se

considera que a vítima mortal é a que não morre no local, mas sim a caminho do hospital, ou no hospital. Nesse caso, será contabilizada na coluna dos mortos/falecidos. Penso que o leitor me acompanha neste raciocínio…

Quanto aos «operacionais», será que o objectivo é designar indistintamente bombeiros e bombeiras, acompanhando os anglófonos que há muito deixaram de falar em *policeman* e *policewoman* para falar em *police officer*? Uma nota positiva nestas trágicas maneiras de dizer — os profissionais dos incêndios, em Portugal, incluindo o pessoal da protecção civil, não comete estes erros, o que me leva a concluir que o mal estará na formação das gentes da tal comunicação. De tão pateta que é, não posso deixar de referir o que um repórter de um canal televisivo afirmou sobre um determinado incêndio: «o incêndio está a deflagrar desde 4.ª-feira, em Oleiros». Por isso mesmo, caro leitor, parece-me útil clarificar o sentido de alguns vocábulos mal empregues por jornalistas e comentadores televisivos: «explosão», «deflagração» e «detonação».

Vejamos. Basicamente, e de uma forma mais ou menos simples, uma explosão é caracterizada por deslocação de matéria em combustão a que se dá o nome de *frente de chama* e deslocação de gás sob pressão libertado que se designa por *onda de choque*. Podemos distinguir dois tipos de explosão: a *deflagração* e a *detonação*. No primeiro caso, a frente de chama desloca-se a uma velocidade subsónica, ou seja, mais lentamente do que a onda de choque; dito de outro modo, a frente de chama desloca-se a uma velocidade inferior à velocidade do som (340 m/s), no ar, à temperatura de 15°. Por exemplo, a chama num queimador de gás é um exemplo de deflagração controlada, já que a velocidade da frente de chama é igual à velocidade de admissão do gás inflamável. Porém, a deflagração pode transformar-se em *detonação* se a frente de chama se deslocar a uma velocidade supersónica, logo superior a 340 m/s, ou seja, superior à velocidade da onda de choque. Não teria cabimento aqui espraiar-me sobre estes conceitos, o objectivo sendo, tão-só, aclarar as diferenças entre os vocábulos, e não maçar o leitor com perorações sobre explosivos.

Chaves certas, completas, e ordenadas

Na esteira deste despautério, entra a lógica, ou talvez fique à porta. Também aqui, o leitor já adivinhou onde quero chegar. Com frequência, na *Antena 1*, estação radiofónica portuguesa que emite em modulação de frequência (já agora, não digam frequência modulada, porque a frequência não está previamente modulada; aliás, atentemos no inglês, ou no francês. *Frequency modulation* que em português significa exactamente «modulação de frequência»; não estamos perante uma «*modulated frequency*»; do mesmo modo, em francês, temos *modulation de fréquence*), mas dizia eu, oiço, por ocasião dos sorteios do Euromilhões, o jornalista de serviço anunciar: «…Euromilhões — vale a pena conferir a chave completa, ou a chave certa do concurso…». Pensemos um pouquito. A chave, por definição, é completa e certa, ou não seria chave; além disso, para ganhar o Euromilhões não é necessário que os números saiam segundo uma determinada ordem sequencial, ascendente ou descendente, por exemplo, ou a ordem de saída da tômbola. Os cinco números e as duas estrelas saem e esses sete números constituem a chave e a ordenação não é condição para ganhar. Se há uma chave, essa chave é necessariamente completa e certa; ordenada ou não, não importa; o que importa é que aqueles números foram sorteados, independentemente da sequência. Estamos, também neste caso, perante uma redundância.

Duas expressões jurídicas inglesas erradamente traduzidas

Caro leitor, mudemos radicalmente de assunto. Não raras vezes, direi mesmo muitíssimas vezes, me deparo com traduções portuguesas de textos jurídicos ingleses em que o sentido de algumas expressões é mal vertido, penso que por desconhecimento do direito, mas também por dificuldade de redacção em português. Uma das expressões muito comuns em contratos, por exemplo, é *terms and conditions* vertido, atrevo-me a dizer, em 100% dos casos, até por juristas como «termos e condições». Vejamos então o verdadeiro sentido da expressão: No direito britânico *conditions* são «*major terms of a contract*», ou seja, as condições fundamentais; já *warranty*, por exemplo, é um «*minor term of a contract*», donde se conclui que *terms* é um vocábulo genérico. Assim, no que se designa por *common*

law, «*terms and conditions*» abrange todas as condições possíveis num contrato, pelo que basta empregarmos em português a expressão «condições», «cláusulas», ou «condições gerais». Ao vertermos por «termos e condições», o que sucede, na grande maioria dos casos, é pensar que *terms* significa termos (na acepção que em português corresponde a «redacção». Atentemos na frase: «*Upon and subject to the terms of the agreement*» (sob reserva das condições do contrato/condições contratuais, e não sob reserva dos termos [redacção] do contrato). Já agora, *term of a contract* também pode fazer referência à «duração do contrato» Outra expressão quantas vezes mal compreendida e vertida para português é *termination*; neste caso, surge frequentemente a confusão entre «rescisão», «resolução», «denúncia». Quanto à rescisão, trata-se de pôr termo, por exemplo, a um contrato, para o tempo futuro, ou seja, sem efeito retroactivo; a resolução é uma anulação, por exemplo, de uma venda, que retroage; na denúncia, estamos perante uma decisão unilateral de uma das partes. Por vezes, o jurista inglês especifica e escreve *earlier termination* que facilita a tradução, existindo, ainda, a opção «expiração do contrato», e tudo dependerá do contexto. Caro leitor, deixo-lhe mais uma expressão que causa sempre dificuldades ao ser vertida para português, creio que por razões que se prendem com a redacção. *From time to time* — não faltam opções: «periodicamente», «pontualmente», «em qualquer momento», «em caso de necessidade/havendo necessidade»; «de tempos a tempos» (esta última sendo o sentido primeiro, em textos não jurídicos).

Ainda e sempre a «polémica escusada» em «a folha» n.º 53

Não podia deixar de alinhavar algumas — brevíssimas — notas acerca da contribuição do Jorge Madeira Mendes(1), desde logo pelo título. Escusada, porquê? Porque se esgotaram os argumentos de uns e outros, ou só de uns, ou só de outros? Escusada porque, e é o sentimento que perpassa, os que são a favor do Acordo Ortográfico (AO) se julgam detentores da verdade absoluta, da razão última, do derradeiro argumento, do raciocínio esclarecido, da verdade científica, dogmática, inabalável, infalível? Escusada porque os detractores pura e simplesmente pensam como uns tontos; como se emitissem radiação divergente que tem de passar por um colimador para se obter um feixe direccionado. Os docentes da disciplina de língua portuguesa da minha geração, nos liceus, em Portugal, são radicalmente contra o AO, tal como a Sociedade Portuguesa de Autores o é. São tontos? Sem pôr (o AO mantém o diacrítico para distinguir da preposição, mas suprime-o em «pára»; então e a confusão com a preposição?) em causa as razões avançadas por uns e outros, todos com o mérito que se lhes deve pelo conhecimento que espelham, permito-me, como tradutor e utilizador da língua — retrógrado, arrogante, intelectualóide (atributos com que foram catalogados em Lisboa os peticionários contra o Acordo) — relembrar que a língua inglesa com as suas variantes (americana, australiana, britânica, neozelandesa, sul-africana e outras faladas noutras partes do continente africano) não esquecendo as diferentes pronúncias no Reino Unido, mormente algumas que ganham terreno nalgumas regiões e entre algumas classes sociais que extravasam Londres (Estuary English, por exemplo) tem palavras com consoantes mudas que não foram guilhotinadas e os falantes, linguistas e governantes ingleses tão-pouco se preocupam; o catalão, o aragonês, o basco, e as variantes latino-americanas do espanhol não incomodam os falantes nem os linguistas, nem os governantes desses países. Serão tontos? Dos PALOP, Angola, Moçambique, Guiné-Bissau, São Tomé e Príncipe e Timor-Leste ficaram de fora. Serão tontos? O acordo, que não passa disso mesmo, pois não estamos perante uma reforma, mas sim perante um acordo sem verdadeiro estudo científico que o sustente, transformou a escrita fonológica e morfológica numa escrita fonética. Recomendo a leitura do artigo saído no *Público* de 14 de Junho de 2017 da autoria de Francisco Miguel Valada(2), investigador e ex-intérprete de conferência no Tribunal de Justiça das Comunidades Europeias, no Luxemburgo. Quer-me parecer que estamos perante um radicalismo abusivo, na fronteira do mau gosto por fazer imperar uns, e decapitar outros pela sua heresia intelectual. Por fim, uma duvidazita: estamos mesmo perante uma polémica, ou será que se trata de uma celeuma?

A literatura universal explicada em português nas ondas hertzianas

Na *Antena 1*, às 4.as-feiras, pelas 23 horas, Ana Daniela Soares reúne à mesa do estúdio Rita Ferro, Inês Pedrosa e Patrícia Reis, num programa que dá pelo nome de «A páginas tantas…». Parabéns ao programa pela variedade dos autores, simplicidade da linguagem, sinceridade nos comentários. Um

programa como há poucos, muito poucos, e que fazem falta, muita falta. Há ocasiões em que o informalismo da linguagem raia a incorrecção, mas é próprio do discurso falado, descontraído e em que a preocupação são os livros, os autores, as editoras, as correntes literárias, as traduções das obras, as opiniões, e o fundo acaba por ser bem mais importante do que a forma. Faço votos para que o formato se mantenha e, porque não, o registo, permitindo que temas maçudos, em rádio, portanto sem a facilidade da imagem para o ouvinte, sejam convidativos e nos prendam durante quase uma hora.

philippe.gariso@mota-engil.pt

(1) Mendes, J. M., «A polémica escusada» *in* «a folha», n.º 53 — primavera de 2017, http://ec.europa.eu/translation/portuguese/magazine/documents/folha53_pt.pdf.
(2) Miguel, F., «Sobre a "opinião dos linguistas", a arrogância, a ignorância e a continência», *Público*, 14 de junho de 2017, https://www.publico.pt/2017/06/14/culturaipsilon/noticia/sobre-a-opiniao-dos-linguistas-a-arrogancia-a-ignorancia-e-a-continencia-1775401.

Notas sobre povos, línguas e topónimos bantos de Angola

Paulo Correia
Amarílis Pêgo (antiga estagiária)
Direção-Geral da Tradução — Comissão Europeia

Uma das características das línguas bantas prende-se com a utilização de **prefixos** (e não sufixos, como em português) para identificar o nome de um ou mais indivíduos de um povo (etnónimo), o nome da língua (glossónimo) que falam ou o nome do território que habitam (topónimo). No quadro seguinte, apresentam-se alguns exemplos.

raiz	**etnónimo**		**glossónimo**	**topónimo**
	sing. (mu-, mo-, …)	**pl.** (ba-, va-, …)	(se-, xi-, qui-, …)	(Le-, Bo-, …)
soto	mossoto (en: *Mosotho*)	bassotos (en: *Basotho*)	sessoto (en: *Sesotho*)	Lessoto (en: *Lesotho*)
tsuana	motsuana (en: *Motswana*)	batsuanas (en: *Batswana*)	setsuana (en: *Setswana*)	Botsuana (en: *Botswana*)
zulo	unzulo (en: *umZulu*)	amazulos (en: *amaZulu*)	issizulo (en: *isiZulu*)	Cuazulo (en: *kwaZulu*)
congo	mucongo (en: *Mukongo*)	bacongos (en: *Bakongo*)	quicongo (en: *Kikongo*)	
bundo norte	mumbundo (en: *Mubundu*)	ambundos (en: Ambundu)	quimbundo (en: Kimbundu)	
bundo sul	ochimbundo (en: *Ocimbundu*)	ovimbundos (en: *Ovimbundu*)	umbundo (en: Umbundu)	

A utilização desses prefixos nos aportuguesamentos é pouco frequente, restringindo-se sobretudo a obras da especialidade[1]. Mesmo quando se usam os prefixos bantos de forma coerente, verifica-se geralmente a adaptação às regras de género ou número do português no caso dos etnónimos, com a consequente criação de singulares a partir de plurais (um **ba**congo e não um mucongo) e de plurais híbridos (dois **ba**congo**s** e não dois bacongo). Assim:

- os **bas**soto**s**, o povo que fala **ses**soto; o **Les**soto[2], a terra dos bassotos
- os **bat**suana**s**, o povo que fala **set**suana; o **Bo**tsuana[3], a terra dos batsuanas
- os **ama**zulos[4], o povo que fala **issi**zulo; o **Cua**zulo[5], a terra dos amazulos

No caso angolano, os vocabulários e dicionários em língua portuguesa apenas utilizam esta abordagem exigente num número muito limitado de casos, como é o dos três maiores grupos etnolinguísticos, todos situados no noroeste e centro de Angola, aqueles com que o contacto é mais antigo.

- os **ba**congo**s**, o povo que fala **qui**congo
- os **a**mbundos, o povo que fala **qui**mbundo
- os **ovi**mbundos, o povo que fala **u**mbundo

A abordagem mais cautelosa limita-se à utilização exclusiva da raiz, tanto para o glossónimo como para o etnónimo, declinando-a de acordo com as regras do português. É esta a abordagem adotada pelos vocabulários e dicionários em língua portuguesa para a generalidade dos grupos etnolinguísticos.

- os sotos, o povo que fala soto; o Lessoto, a terra dos sotos
- os tsuanas, o povo que fala tsuana; o Botsuana, a terra dos tsuanas
- os zulos, o povo que fala zulo; o Cuazulo, a terra dos zulos

É, no entanto, frequente encontrar textos com uma utilização algo incoerente dos prefixos bantos[(6)], que, conjugada com a variabilidade ortográfica dos bantoísmos, pode conduzir a uma certa dispersão terminológica. Por vezes, os dicionários normalizam estes «desvios». Tal é o caso do termo quimbundo, registado como língua, mas também como etnónimo.

Povos e línguas nacionais de Angola

O povo angolano é composto por vários grupos etnolinguísticos[(7)], alguns deles vivendo em áreas que vão para além das fronteiras da República de Angola. É o caso dos bacongos, a norte, mas também dos chócues (ou quiocos), a nordeste, ou dos cuanhamas, a sul.

Para além da língua oficial (o português), em Angola falam-se várias **línguas nacionais** (com as suas variantes internas e variantes-padrão), maioritariamente do grupo banto, mas também de grupos pré-bantos (coissã e vátua[(8)]). Em 1979, foi criado o Instituto Nacional de Línguas para estabelecer normas ortográficas e estudar e investigar os aspetos linguísticos de algumas dessas línguas. Em 1987, para ajudar a normalizar a grafia das línguas nacionais, foram aprovados, a título experimental, os alfabetos de seis línguas nacionais.

> Resolução n.º 3/87, de 23 de maio, do Conselho de Ministros
> Artigo 1.º São aprovados a título experimental os Alfabetos das Línguas: «Kikongo», «Kimbundú», «Cokwé», «Umbundu», «Mbunda», e «Oxikwanyama» e as respetivas Regras de Transcrição, em anexo que fazem parte do presente diploma.[(9)]

No quadro seguinte apresentam-se algumas línguas bantas nacionais citadas em fontes oficiais angolanas e os etnónimos correspondentes. Os nomes em português são os registados no Vocabulário Ortográfico Comum da Língua Portuguesa (VOC)[(10)] ou em dicionários de língua portuguesa.

ISO 639-3	**glossónimo etnónimo**			**%**[(11)]	**outros países (ISO)**	**IATE**
	pt	**língua nacional**	**outras var.**			
kng	quicongo[(12)] (congo, conguês)	*Kikongo*[a,b,c] *Ukongo*[c]	*Kongo* *Kicongo* *Kikoongo*	8,24	CG CD	1451582
	bacongo(s)	*Bakongo*				3546806
kmb	quimbundo (bundo do norte)	*Kimbundu*[a,b,c]	*Kimbundo*	7,82		3572162
	ambundo(s)	*Mbundu*	*Ambundu* *Mbundo*			918075

umb	umbundo (bundo do sul)	*Umbundu*[a,b,c]		22,96		3572167
	ovimbundo(s)	*Ovimbundu*				918077
cjk	chócue (chichocué)	*Cokwe*[a,b] *Chokwe*[c] *Kioko*[c]	*Tchócue* *Tchokwe* *Chichokwe* *Côkwe* *Xokwe* *Tchokwe* *Kyoko*	6,54	CD ZM	3572165
	chócue(s) (quiocos) (tuchócues)	*Cokwe*	*Chokwe* *Tucokwe*			3546809
kua	cuanhama[13] (oxicuanhama)	*Oxiwambo*[a] *Oxikwanyama*[b] *Kwanhama*[c]	*Oxicuanhama* *Oxikwanyama* *Oshikwanyama* *Kwanyama*	2,26	NA	3572168
	cuanhama(s) (ovacuanhamas)	*Ovakwanyama*	*Aawambo* *Ambo* *Aawambo* *Ovawambo*			3572192
mck	bunda (chimbunda)	*Mbunda*[a,b]	*Mbuunda* *Chimbuunda*	—[14]	ZM	3572197
	bundo(s) (vambundos)	*Mbunda*	*Vambunda*			3572201
nba	ganguela	*Ngangela*[a] *Nganguela*[c]	*Nyemba* *Luchazi*	3,11		3572191
	ganguela(s) (vanganguelas)	*Vangangela*	*Vanganguela*			3546810
nyk	nhaneca[15] (olunhaneca)	*Olunyaneka*[a] *Nhaneca*[c]	*Nyaneka* *Lunyaneka*	3,42		3573821
	nhaneca(s)[16] (ovanhanecas)	*Ovanyaneka*				3546812
lue	luvale (chiluvale)	*Luvale*[a,c]	*Chiluvale*	1,04	ZM	3573858
	luvale(s)	*Luvale*				3573859
her	herero (oxierero)	*Helelo*[a]	*Oshihelelo* *Otjiherero*		NA BW	3573857
	herero(s) (ovaereros)	*Ovahelelo*				3546815

Fontes angolanas:
a — Projeto de lei sobre o estatuto das línguas nacionais, Ministério da Cultura da República de Angola, agosto de 2011, http://www.mincult.gov.ao/download.aspx?id=836&tipo=publicacao.
b — Resolução n.º 3/87 que aprova a título experimental, os Alfabetos de 6 Línguas Nacionais, *Diário da República*, 1.ª Série, n.º 41, 23 de maio de 1987, http://www.embaixadadeangola.org/cultura/linguas/set_lnac.html.
c — Instituto Nacional de Estatística da República de Angola, *Angola — Recenseamento Geral da População e Habitação — 2014*, https://andine.ine.gov.ao/nada4/index.php/catalog/14.

Toponímia plurilingue

> Cuando Cubango? Kuando Kubango? Kwando Kubango? Kuwando Kubango?
> Cuanza? Quanza? Kuanza? Kwanza? Kuwanza?
> Luanda? Loanda? Lwanda? Luwanda?
> Benguela? Bengela? Mbengela?

Observa-se em fontes angolanas em língua portuguesa, mesmos oficiais, a coexistência de toponímia com grafias quer segundo as regras do português quer segundo as regras ortográficas das línguas bantas locais, ou ainda segundo formas híbridas banto-português, refletindo diferentes influências e

práticas. A recente Lei de Bases da Toponímia, parece ter como objetivo enquadrar melhor situações como estas:

> Lei n.º 14/16, de 12 de setembro, da Assembleia Nacional
> Lei de Bases da Toponímia, que estabelece as bases para a definição e disciplina da toponímia ao nível nacional e local, bem como as regras e procedimentos para efeitos de atribuição de números de polícia.
> (…)
> Artigo 7.º
> (Regras de Grafia dos Topónimos)
> 1. Os topónimos são escritos em língua portuguesa, seguindo a grafia latina.
> 2. Os topónimos, nas demais línguas de Angola, são escritos em conformidade com as regras de grafia da língua, devendo ser certificados pelo Instituto de Línguas Nacionais.(17)

Como nas regiões de Bruxelas ou de Miranda do Douro, estará Angola a avançar para placas toponímicas bilingues?

No quadro seguinte, apresentam-se as atuais 18 províncias angolanas, de norte para sul e de oeste para leste e as respetivas capitais e gentílicos provinciais. Além da ortografia em português, indica-se para as províncias e capitais uma ortografia banta. Para o nome das províncias em português utilizou-se a ortografia empregada pelo Instituto Nacional de Estatística (INE) de Angola, por exemplo, no inquérito de indicadores múltiplos e de saúde de 2015(18). Excetua-se o caso do Uíje (por aproximação à ortografia banta, utilizou-se aí «j» em vez de «g»), em que foram utilizadas outras fontes(19). Foram criadas fichas IATE em português, quimbundo(20) e inglês. Para os nomes das capitais provinciais optou-se pela aférese (supressão de sons) no caso das consoantes iniciais com pré-nasalação. Assim, Banza (para Mbanza), Benguela (para Mbengela) e Dalatando (para Ndalatandu). Para os gentílicos provinciais evitaram-se etnónimos, pois numa mesma província podem viver vários povos.

ISO 3166	**província(21)**		**capital**		**gentílico provincial**	**IATE**
AO-	**pt**	**língua nacional**	**pt**	**língua nacional(22)**		
CAB	Cabinda	Kabinda	Cabinda	Kabinda	cabindense(s)	2229256
ZAI	Zaire	Nzadi	Banza Congo(23)	Mbanza Kongo	zairense(s)	3572096
UIG	Uíje	Wiji	Uíje(24)	Wiji	uijense(s)	3572097
LUA	Luanda	Lwanda	Luanda	Lwanda	luandense(s)(25)	3572098
BGO	Bengo	Mbengu	Caxito(26)	Kaxitu	benguense(s)	3572101
CNO	Cuanza Norte	Kwanza-Norte	Dalatando(27)	Ndalatandu	norte-cuanzense(s)	3572122
MAL	Malanje	Malanji	Malanje	Malanji	malanjino(s)/a(s)(28)	3572095
LNO	Lunda Norte	Lunda-Norte	Dundo(29)	Dundu	norte-lundense(s)	3572125
CUS	Cuanza Sul	Kwanza-Sul	Sumbe(30)	Sumbi	sul-cuanzense(s)	3572136
LSU	Lunda Sul	Lunda-Sul	Saurimo(31)	Saurimu	sul-lundense(s)	3572134
BGU	Benguela	Mbengela	Benguela	Mbengela	benguelense(s)	3572137
HUA	Huambo	Wambu	Huambo(32)	Wambu	huambense(s)33	3572144
BIE	Bié	Viye	Cuíto(34)	Kwitu	bieno/a(s)(35)	3572146
MOX	Moxico	Moxiku	Luena(36)	Lwena	moxiquense(s)	3573737
NAM	Namibe	Namibi	Moçâmedes(37)	Musamedis(38)	namibense(s)	3572147
HUI	Huíla	Wila	Lubango(39)	Lubangu	huilense(s)	3572102
CNN	Cunene	Kuneni	Onjiva(40)	Ondjiva	cunenense(s)	3572148
CCU	Cuando Cubango	Kwandu Kuvangu	Menongue(41)	Menongi	cuando-cubanguense(s)	3572149

Paulo.Correia@ec.europa.eu
amar.pego@gmail.com

[1] Cf. os glossónimos sessoto, setsuana ou issizulo em Petter, M. (org.), *Introdução à Linguística Africana*, Editora Contexto, ISBN 978-85-7244-934-2.
[2] Até 1966, Basutolândia, ou melhor, Bassotolândia — Terra dos Bassotos.
[3] Até 1966, Bechuanalândia, ou melhor, Batsuanalândia — Terra dos Batsuanas.
[4] O *Aulete Digital* define este termo como «cafres da costa oriental africana», http://www.aulete.com.br/amazulos.
[5] Cuazulo-Natal é uma das entidades regionais da África do Sul (antiga província do Natal).
[6] Exemplo: «A população de Angola é maioritariamente de origem Bantu. Dentre estes destacam-se os Ovimbundus (grupo com maior expressão), os Kimbundus e os Bakongos.», Guia Turístico de Angola, *Angola: População*, http://www.guiaturisticoangola.co.ao/index.aspx?menuid=5&lang=P.
[7] *Il faut se mettre en garde contre le sens caché de certaines expressions. On ne choque pas en employant le terme d'« ethnie zoulou » alors que « ethnie portugaise » étonne. Ce choix de terme reproduit un décalage entre « civilisé » et « sauvage », héritage de la période colonisatrice.*, Cahen, M., *Retour de l'ethnicité ?*, Centre Culturel Hâ 32, http://ha32.org/spip/spip.php?article71.
Não se falará de etnias, mas de povos angolanos (a etnia alemã na Bélgica?) e não se usarão formas invariáveis em género e número ou com maiúsculas iniciais para os respetivos nomes e adjetivos (os Flamengo(s) e os Valão(es) da Bélgica?).
Se se compararem, por exemplo, os artigos de apoio da Infopédia sobre dois Estados plurinacionais como Angola e Bélgica, verifica-se que as palavras etnia e étnico são usados para o primeiro país, mas não são para o segundo.
Cf. «A principal etnia é a Ovimbundo, que constitui cerca de 2/5 da população total do país e que fala umbundo.», Porto Editora, *Infopédia: Língua Portuguesa*, «Artigos de apoio: Angola», https://www.infopedia.pt/$angola, e «Os flamengos representam cerca de 58% da população e os valõe**s** 31%», Porto Editora, *Infopédia: Língua Portuguesa*, «Artigos de apoio: Bélgica», https://www.infopedia.pt/$belgica.
Ver também, a este propósito, Correia, P., «Etnónimos, uma categoria gramatical à parte?» *in* «a folha», n.º 40 — outono de 2012, http://ec.europa.eu/translation/portuguese/magazine/documents/folha40_pt.pdf.
[8] Não confundir os vátuas do sul de Angola com os vátuas do sul de Moçambique. Estes últimos são um povo banto.
[9] Resolução n.º 3/87 que aprova a título experimental, os Alfabetos de 6 Línguas Nacionais, *Diário da República*, 1.ª Série, n.º 41, 23 de maio de 1987, http://www.embaixadadeangola.org/cultura/linguas/set_lnac.html.
[10] Instituto Internacional da Língua Portuguesa, *Vocabulário Ortográfico Comum da Língua Portuguesa*, http://voc.iilp.cplp.org/.
[11] Instituto Nacional de Estatística da República de Angola, *Angola — Recenseamento Geral da População e Habitação — 2014*, https://andine.ine.gov.ao/nada4/index.php/catalog/14.
[12] Algumas fontes apresentam também, separadamente, o **fiote**, variante do quicongo falada em Cabinda.
[13] Com o xindongo (ndo) formam o ovambo.
[14] Não aparece no *Recenseamento Geral da População e Habitação - 2014* do Instituto Nacional de Estatística da República de Angola. Foi substituída como língua nacional pela língua ganguela, cf. Wikipedia, *Mbunda*, https://ca.wikipedia.org/wiki/Mbunda.
[15] O VOC regista igualmente lunhaneca.
[16] O VOC regista igualmente munhaneca(s).
[17] Lei n.º 13/16 (Lei de Bases da Organização Administrativa do Território) que estabelece as bases para a organização do território da República de Angola, para fins político-administrativos e a designação, criação, classificação e progressão das unidades urbanas e outros aglomerados populacionais, *Diário da República*, 1.ª Série, n.º 155, 12 de setembro de 2016.
[18] Instituto Nacional de Estatística da República de Angola, *Inquérito de Indicadores Múltiplos e de Saúde (IIMS) – 2015-2016: Relatório de Indicadores Básicos*,
http://www.ine.gov.ao/xportal/xmain?xpid=ine&xpgid=publications_detail&publications_detail_qry=BOUI=42933913
Confrontar igualmente:

- Governo da República de Angola, *Governos Provinciais*, http://www.angola.gov.ao/GovernosProvinciais.aspx
- Embaixada da República de Angola em Portugal, *Províncias*, http://www.embaixadadeangola.pt/726-2/

[19] «Miguel Cutoca disse que o Ministério da Justiça quer evitar que os cidadãos residentes no interior da província percorram longas distâncias até à **cidade do Uíje** para obterem o Bilhete de Identidade ou o Registo Criminal. O delegado da Justiça disse que nos municípios da **província do Uíje** já funcionam os serviços de Registo Civil e Notariado e nesta altura o Ministério da Justiça trabalha, em colaboração com o Governo Provincial, na construção de infraestruturas para as Repartições Municipais de Identificação Civil e Criminal», Governo da República de Angola, *Notícias*, «Acesso à justiça em debate», 8 de dezembro de 2014, http://www.cidadao.gov.ao/VerNoticia.aspx?id=25277.
[20] Para as províncias de Luanda, Bengo, Cuanza Norte, Malanje, onde o quimbundo é a língua nacional maioritária, também para as províncias do Uíje, Cuanza Sul, Lunda Norte e Lunda Sul, onde também é falada.
[21] Até 1975, as subdivisões administrativas eram designadas distritos. Cf. Lei n.º 2066, de 27 de junho de 1953, que promulga a Lei Orgânica do Ultramar Português, *Diário do Governo*, I série, n.º 135, de 27 de junho de 1953, https://dre.pt/application/conteudo/640659.
Em Angola, o termo distrito refere-se atualmente a subdivisões dos municípios.
[22] Nalguns casos, escrita com ortografia banta de nomes portugueses ou aportuguesados.
[23] Etimologicamente, Cidade do Congo. Até 1975, São Salvador do Congo.
[24] Até 1975, Carmona.
[25] Cf. caluanda — do quimbundo *mukwa lwanda*, natural ou habitante de Luanda (cidade).
[26] Etimologicamente, Kaxitu é diminutivo de *xitu* (carne), significando pedaço de carne.
[27] Até 1975, Vila Salazar.
[28] malanjinho / malanjense
[29] Até 1975, Portugália.

[30] Até 1975, Novo Redondo.
[31] Até 1975, Henrique de Carvalho.
[32] Até 1975, Nova Lisboa.
[33] Cf. huambo — do povo huambo.
[34] Até 1975, Silva Porto.
[35] Porto Editora, *Infopédia: Língua Portuguesa*, «bieno», https://www.infopedia.pt/dicionarios/lingua-portuguesa/bieno.
[36] Até 1975, Luso.
[37] «A localidade foi batizada em 1485 com o nome de Tchitoto Cho Patua, depois Mussungo Bitoto. Em 1775, chamou-se Angra dos Negros e, mais tarde, Pinheiro Furtado, de Baía de Mossâmedes, na altura com "ss", em homenagem a José de Almeida Vasconcelos de Oliveira Soveral de Carvalho, Barão de Mossâmedes [povoação da freguesia de São Miguel de Mato, Vouzela, Viseu, Portugal]. A cidade foi fundada em 1840 e passou, em 14 de novembro de 1956, a escrever-se Moçâmedes com "c" de cedilha [indicando a origem árabe do topónimo], até 1975, quando passou a chamar-se Namibe. Em 27 de junho do corrente ano [2016], o nome voltou a ser alterado por decreto-lei para cidade de Moçâmedes.», Sousa, M., «Capital do Namibe volta a ser Moçâmedes», *Jornal de Angola*, 4 de agosto de 2016, http://jornaldeangola.sapo.ao/sociedade/capital_do_namibe_volta_a_ser_mocamedes.
[38] Possível transcrição.
[39] Até 1975, Sá da Bandeira.
[40] Até 1975, Vila Pereira de Eça.
[41] Até 1975, Serpa Pinto.

Letónia — ficha de país

Juta Preisa
Secretariado-Geral — Conselho da União Europeia
Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

Nesta ficha de país reúne-se informação terminológica relativa à Letónia que se encontra dispersa por vários documentos normativos ou de referência das instituições europeias.

Apresenta-se em anexo a esta ficha uma tabela com o alfabeto letão, os símbolos fonéticos correspondentes às letras e transliterações aproximadas em português.

REPÚBLICA DA LETÓNIA (IATE: 861101)

CAPITAL: Riga
GENTÍLICO/ADJETIVO: letão/letã (plural: letões/letãs)
MOEDA: euro
SUBDIVISÃO: cent[1]

Principais cidades: Riga, Daugavpils, Liepāja
Rios: Daugava
Serras: Gaiziņkalns

Subdivisões administrativas

#	letão (Eurostat)	português	inglês	IATE
9	Republikas pilsēta	cidade da República	republic city	3556134
110	novads	município	municipality	3522567

Fonte: Eurostat, *Nomenclature of territorial units for statistics: National Structures (EU),* http://ec.europa.eu/eurostat/web/nuts/national-structures-eu

Regiões

NUTS	LATVIJA	LETÓNIA	LATVIA	IATE
LV003	Kurzeme	Curlândia	Kurzeme	3573886
LV005	Latgale	Latgália(2)	Latgale	3574023
LV006	Rīga	Riga	Rīga	1891480
LV007	Pierīga	Pieriga(3)	Pierīga	—
LV008	Vidzeme	Vidzeme(4)	Vidzeme	3573887
LV009	Zemgale	Semigália(5)	Zemgale	3574025

Fonte: Serviço das Publicações, *Código de Redação Interinstitucional*: Anexo 10 — Lista das Regiões, http://publications.europa.eu/code/pt/pt-5001000.htm
Regulamento (CE) n.º 1059/2003 do Parlamento Europeu e do Conselho, de 26 de maio de 2003, relativo à instituição de uma Nomenclatura Comum das Unidades Territoriais Estatísticas (NUTS) (versão consolidada de 2016), http://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:02003R1059-20161219.

Órgãos judiciais

#	letão	português	inglês	IATE
27	administratīvā rajona tiesa	tribunal administrativo de comarca	administrative district court	3545842
4	administratīvā apgabaltiesa	tribunal administrativo regional	administrative regional court	3545841
34	priekšpilsētas tiesa	tribunal metropolitano	city district court	3522536
	rajona tiesa	tribunal de comarca(6)	district court	3522537
6	apgabaltiesa	tribunal regional	regional court	3522535
1	Augstākā tiesa	Supremo Tribunal	Supreme Court	3522533
1	Satversmes tiesa	Tribunal Constitucional	Constitutional Court	3522532

Portal Europeu da Justiça. *Sistemas judiciais nos Estados-Membros — Letónia*, https://e-justice.europa.eu/content_judicial_systems_in_member_states-16-lv-pt.do?member=1

Juta.Preisa@consilium.europa.eu
Paulo.Correia@ec.europa.eu

Anexo: Alfabeto letão

O alfabeto letão é derivado do alfabeto latino, incluindo onze carateres suplementares formados por adição de diacríticos:

A, Ā, B, C, Č, D, E, Ē, F, G, Ģ, H, I, Ī, J, K, Ķ, L, Ļ, M, N, Ņ, O, P, R, S, Š, T, U, Ū, V, Z, Ž

As letras Q, W, X e Y são utilizadas apenas em palavras estrangeiras, sobretudo nomes, quando são escritos na língua original. Os dígrafos DZ, DŽ, LL e ĻĻ, não fazem parte do alfabeto.

Existem ainda os ditongos: AI, IE, EI, EU, O (/uɑ̯/), AU, UI.

letra letã	fonética (AFI)	equivalente português	nome letão	«transliteração»
A a	/ɑ/	a curto	***a**kmens* (pedra)	akmens
Ā ā	/ɑː/	a longo	***ā**bele* (macieira)	aabele
B b	/b/	b (em *b*ola)	***b**umba* (bola)	bumba
C c	/t͡s̪/	ts (em tsunâmi)	***c**ukurs* (açúcar)	tsukurs
Č č	/tʃ/	tch (em *Tch*aikovsky)	***č**ačača* (chachachá)	tchatchatcha
D d	/d̪/	d (em *d*ente)	***d**rupas* (ruínas)	drupass
E e	/e/ /æ/	e (em p*e*lo) e (em sereia)	***e**gle* (abeto) ***s**ekas* (consequências)	egle sekass
Ē ē	/eː/ /æː/	e longo e longo lisboeta (em leite)	***ē**zelis* (burro) ***ē**na* (sombra)	eezeliss eena
F f	/f/	f (em *f*aca)	***f**azāns* (faisão)	fazaans
G g	/g/	g (em *g*ato)	***g**aiss* (ar)	gaiss

Ģ ģ	/ɟ/	d francês (em *d*ieu)	***ģ**imene* (família)	dimene
H h	/x/	r lisboeta (em *r*osa)	***h**okejs* (hóquei)	roqueiss
I i	/i/	i (em v*i*da)	***i**staba* (quarto)	istaba
Ī ī	/iː/	i longo	***ī**re* (aluguer)	iire
J j	/j/	i (em cu*i*dado)	***j**umts* (telhado)	iumts
K k	/k/	c (em *c*asa)	***k**aste* (caixa)	kaste
Ķ ķ	/c/	t francês (em *t*iare)	***ķ**irbis* (abóbora)	tirbiss
L l	/l/	l (em *l*ua)	***l**ikums* (lei)	likums
Ļ ļ	/ʎ/	lh (em i*lh*a)	***ļ**aut* (deixar)	lhaut
M m	/m/	m (em *m*orte)	***m**aize* (pão)	maize
N n	/n̪/ /ŋ/	n (em *n*ada) n (em ma*n*ga)	***n**auda* (dinheiro) *ru**n**ga* (taco)	nauda runga
Ņ ņ	/ɲ/	nh (em so*nh*o)	***ņ**audēt* (miar)	nhaudeet
O o	/u̯a/ /o/ /oː/	ua (em q*ua*se) o (em m*o*rto) o longo	***o**ga* (baga) ***o**āze* (oásis) ***o**da* (ode)	uaga oaaze ooda
P p	/p/	p (em *p*ato)	***p**ils* (castelo; burgo)	pils
R r	/r/	r (em ca*r*o)	**R**īga (Riga)	Riiga
S s	/s̪/	s (em *s*aia)	***s**aule* (sol)	saule
Š š	/ʃ/	x (em *x*á)	***š**alle* (xaile)	xal-le
T t	/t̪/	t (em *t*udo)	***t**ētis* (pai)	teetiss
U u	/u/	u (em per*u*)	***u**guns* (fogo)	uguns
Ū ū	/uː/	u longo	***ū**dens* (água)	uudens
V v	/v/	v (em *v*aca)	***v**alsts* (país)	valsts
Z z	/z̪/	z (em *z*ebra)	***z**eme* (terra)	zeme
Ž ž	/ʒ/	j (em *j*eito)	***ž**āvēt* (secar)	jaaveet
LL ll	/lː/	l longo	***l**elle* (boneca)	lel-le
ĻĻ ļļ	/ʎː/	lh longo	*ba**ļļ**a* (tina de madeira)	balh-lha
DZ dz	/d̪͡z̪/	dz (em *dz*eta)	*dzenis*(7) (pica-pau)	dzeniss
DŽ dž	/d͡ʒ/	dj (em *dj*embê)	***dž**insi* (calças de ganga)	djinsi

(1) Forma obrigatória nos atos da UE e a preferir nos demais textos da UE); cêntimo (variante nacional de uso corrente em Portugal e que pode ser utilizada noutro tipo de textos).

(2) A *Lello Universal* refere Letgália. Cf. Rocha, C., «Os gentílicos de Latgália, Selónia, Semigália e Vidzeme (Letónia)», *Ciberdúvidas*, 1 de fevereiro de 2013, https://ciberduvidas.iscte-iul.pt/consultorio/perguntas/os-gentilicos-de-latgalia-selonia-semigalia-e-vidzeme-letonia/32117.

(3) Significa, literalmente, Circunriga (à volta de Riga).

(4) Corresponde à parte centro-meridional da Livónia histórica, pelo que por vezes também é designada por Livónia. Vidzeme significa, literalmente, Terra do Meio (*zeme* = terra).
Cf. Rocha, C., «Os gentílicos de Latgália, Selónia, Semigália e Vidzeme (Letónia)», *Ciberdúvidas*, 1 de fevereiro de 2013, https://ciberduvidas.iscte-iul.pt/consultorio/perguntas/os-gentilicos-de-latgalia-selonia-semigalia-e-vidzeme-letonia/32117.

(5) A *Lello Universal* refere Semigola. Cf. Rocha, C., «Os gentílicos de Latgália, Selónia, Semigália e Vidzeme (Letónia)», *Ciberdúvidas*, 1 de fevereiro de 2013, https://ciberduvidas.iscte-iul.pt/consultorio/perguntas/os-gentilicos-de-latgalia-selonia-semigalia-e-vidzeme-letonia/32117.

(6) Tribunal de comarca é o termo normalizado que tem sido utilizado para tribunais de primeira instância nas diferentes fichas de país publicadas n'«a folha».

(7) A palavra Vidzeme é uma exceção à regra pois é uma palavra composta. O «d» e o «z» pertencem, assim, a duas palavras distintas — *vi**d*** (de *vi**d**us* (meio)) + **z**eme (terra) — e não podem ser fundidos no som «dz». Esta palavra pronuncia-se Vid-zeme.

Redação: Paulo Correia (Comissão); Valdemar Ferreira (PE); Manuel Leal (Conselho da UE); Victor Macedo (CESE-CR); António Raúl Reis (Serviço das Publicações)
Grupo de apoio: Ana Luísa Faria (Conselho da UE); Hilário Leal Fontes (Comissão); Susana Gonçalves (Comissão); Ana Lorenzo Garrido (Comissão); Joana Seixas (CESE-CR)
Paginação: Susana Gonçalves (Comissão)
Envio de correspondência: dgt-folha@ec.europa.eu

Edição impressa: oficinas gráficas do Serviço de Infraestruturas e Logística — Bruxelas (Comissão)
Edição eletrónica: sítio Web da Direção-Geral da Tradução da Comissão Europeia no portal da União Europeia — http://ec.europa.eu/translation/portuguese/magazine

Centro de Informação Europeia Jacques Delors
DGAE/Ministério dos Negócios Estrangeiros

A coleção completa d'«a folha» está disponível no catálogo em linha da Biblioteca Jacques Delors
https://infoeuropa.eurocid.pt/pesquisar/wti=(a+folha)+AND+(wfmt=se+OR+wfmt=an)/catalogo=bibliografico

«a folha» ISSN 1830-7809

ISSN 1830-780-9
9 771830 780004